ISRAEL JOURNAL OF MATHEMATICS **180** (2010), 1–41
DOI: 10.1007/s11856-010-0092-z

CHARACTERIZATION OF WEAK CONVERGENCE OF BIRKHOFF SUMS FOR GIBBS–MARKOV MAPS

BY

Sébastien Gouëzel

IRMAR, CNRS UMR 6625, Université de Rennes 1
35042 Rennes, France
e-mail: sebastien.gouezel@univ-rennes1.fr

ABSTRACT

We investigate limit theorems for Birkhoff sums of locally Hölder functions under the iteration of Gibbs–Markov maps. Aaronson and Denker have given sufficient conditions to have limit theorems in this setting. We show that these conditions are also necessary: there is no exotic limit theorem for Gibbs–Markov maps. Our proofs, valid under very weak regularity assumptions, involve weak perturbation theory and interpolation spaces. For L^2 observables, we also obtain necessary and sufficient conditions to control the speed of convergence in the central limit theorem.

1. Introduction and results

Let T be a probability preserving transformation on a space X, and let $f : X \to \mathbb{R}$. We are interested in this paper in limit theorems for sequences $(S_n f - A_n)/B_n$, where $S_n f = \sum_{k=0}^{n-1} f \circ T^k$ and A_n, B_n are real numbers with $B_n > 0$. If T is a Gibbs–Markov map and f satisfies a very weak regularity assumption, we will give necessary and sufficient conditions for the convergence in distribution of $(S_n f - A_n)/B_n$ to a nondegenerate random variable. Sufficient conditions for this convergence are already known by the work of Aaronson and Denker [AD01b, AD01a] (under stronger regularity assumptions), and the main point of this article is to show that these conditions are also necessary. We will

Received September 3, 2008

also considerably weaken the regularity assumptions of Aaronson and Denker, by using weak perturbation theory [KL99, Her05].

Finding necessary conditions for limit theorems in dynamical systems has already been considered in [Sar06], but here the author considered only random variables in a controlled class of distributions, while our results apply to all random variables. The paper [Jak93] (see also [DJ89]) gives in a wider setting (the condition (B) in this paper is satisfied for Gibbs–Markov maps) a partial answer to the questions we are considering: if one assumes that $A_n = 0$, then the limiting distribution has to be stable, as in the case of i.i.d. random variables. However, it does not describe for which functions f the convergence $S_n f/B_n \to W$ takes place, nor does it treat the more difficult case $A_n \neq 0$.

At the heart of our argument lies a very precise control on the leading eigenvalue of perturbed transfer operators: if the function f belongs to L^p for $p \in (1, \infty)$, we obtain such a control up to an error term $O(|t|^{p+\epsilon})$ for some $\epsilon > 0$, in Theorem 3.13. This estimate is useful in many different situations: we illustrate it by deriving, in Appendix A, necessary and sufficient conditions for the Berry–Esseen theorem (i.e., estimates on the speed of convergence in the central limit theorem), for L^2 observables satisfying the same weak regularity condition as above.

1.1. The case of i.i.d. random variables. Since our limit theorems will be modeled on corresponding limit theorems for sums of independent identically distributed random variables, let us first describe the classical results in this setting (the statements of this paragraph can be found in [Fel66] or [IL71]).

Definition 1.1: Let X_n be a sequence of random variable. This sequence satisfies a **nondegenerate limit theorem** if there exist $A_n \in \mathbb{R}$ and $B_n > 0$ such that $(X_n - A_n)/B_n$ converges in distribution to a nonconstant random variable.

Definition 1.2: A measurable function $L : \mathbb{R}_+^* \to \mathbb{R}_+^*$ is **slowly varying** if, for any $\lambda > 0$, $L(\lambda x)/L(x) \to 1$ when $x \to +\infty$.

We define three sets of random variables as follows:

- Let $\mathcal{D}_1$ be the set of nonconstant random variables Z whose square is integrable.
- Let $\mathcal{D}_2$ be the set of random variables Z such that the function $L(x) := E(Z^2 1_{|Z| \leq x})$ is unbounded and slowly varying (equivalently,

$P(|Z| > x) = x^{-2}\ell(x)$ for a function ℓ such that $\tilde{L}(x) := 2\int_1^x \frac{\ell(u)}{u}\,\mathrm{d}u$ is unbounded and slowly varying, and in this case L and $\tilde{L}$ are equivalent at $+\infty$).

- Finally, let $\mathcal{D}_3$ be the set of random variables Z such that there exist $p \in (0,2)$, a slowly varying function L and $c_1, c_2 \geq 0$ with $c_1 + c_2 = 1$ such that $P(Z > x) = (c_1 + o(1))L(x)x^{-p}$ and $P(Z < -x) = (c_2 + o(1))L(x)x^{-p}$ when $x \to +\infty$.

Let also $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2 \cup \mathcal{D}_3$. The set $\mathcal{D}$ is exactly the set of random variables satisfying nondegenerate limit theorems. We will now describe the norming constants and the limiting distribution in these theorems.

Let $Z \in \mathcal{D}$; let $Z_0, Z_1, \dots$ be i.i.d. random variables with the same distribution as Z. Then

- If $Z \in \mathcal{D}_1$, let $B_n = \sqrt{n}$ and $W = \mathcal{N}(0, E(Z^2) - E(Z)^2)$.
- If $Z \in \mathcal{D}_2$, let $B_n \to \infty$ satisfy $nL(B_n) \sim B_n^2$ and let $W = \mathcal{N}(0,1)$.
- If $Z \in \mathcal{D}_3$, let $B_n \to \infty$ satisfy $nL(B_n) \sim B_n^p$. Define $c = \Gamma(1-p)\cos(p\pi/2)$ if $p \neq 1$ and $c = \pi/2$ if $p = 1$, and $\beta = c_1 - c_2$. Let $\omega(p,t) = \tan(p\pi/2)$ if $p \neq 1$ and $\omega(1,t) = -\frac{2}{\pi}\log|t|$. Let W be the random variable with characteristic function

$$E(e^{\mathbf{i}tW}) = e^{-c|t|^p(1-\mathbf{i}\beta\,\mathrm{sgn}(t)\omega(p,t))}. \tag{1.1}$$

Theorem 1.3: *In all three cases, there exists A_n such that*

$$\frac{\sum_{k=0}^{n-1} Z_k - A_n}{B_n} \to W. \tag{1.2}$$

One can take $A_n = nE(Z)$ if Z is integrable, and $A_n = 0$ if $Z \in \mathcal{D}_3$ with $p < 1$ (if $p = 1$ but Z is not integrable, the value of A_n is more complicated to express, see [AD98]*).*

Moreover, the random variables in $\mathcal{D}$ are the only ones to satisfy such a limit theorem: if a random variable Z is such that the sequence $\sum_{k=0}^{n-1} Z_k$ satisfies a nondegenerate limit theorem, then $Z \in \mathcal{D}$.

The set $\mathcal{D}$ can therefore be described as the set of random variables belonging to a domain of attraction. The limit laws in this theorem are the normal law and the so-called stable laws. The two parts of this theorem are quite different: while the direct implication is quite elementary (it boils down to a computation of characteristic functions), the converse implication, showing that a random variable automatically belongs to $\mathcal{D}$ if it satisfies a nondegenerate limit theorem, is much more involved, and requires the full strength of Lévy–Khinchine theory.

The direct implication of Theorem 1.3 describes one limit theorem for random variables in $\mathcal{D}$, but does not exclude the possibility of other limit theorems (for different centering and scaling sequences). However, the following convergence of types theorem (see, e.g., [Bil95, Theorem 14.2]) ensures that it can only be the case in a trivial way:

THEOREM 1.4: *Let W_n be a sequence of random variables converging in distribution to a nondegenerate random variable W. If, for some $A_n \in \mathbb{R}$ and $B_n > 0$, the sequence $(W_n - A_n)/B_n$ also converges in distribution to a nondegenerate random variable W', then the sequences A_n and B_n converge respectively to real numbers A and B (and W' is equal in distribution to $(W - A)/B$).*

The specific form of the convergence, the norming constants or the limit laws in Theorem 1.3 will not be important to us. Indeed, we will prove in a dynamical setting that Birkhoff sums satisfy a limit theorem if and only if the sums of corresponding i.i.d. random variables also satisfy a limit theorem. Using Theorem 1.3, this will readily imply a complete characterization of functions satisfying a limit theorem — it will not be necessary to look into the details of Theorem 1.3 and the specific form of the domains of attraction, contrary to what is done in [AD01b, AD01a].

1.2. LIMIT THEOREMS FOR GIBBS–MARKOV MAPS. Let (X, d) be a bounded metric space endowed with a probability measure m. A probability preserving map $T : X \to X$ is **Gibbs–Markov** if there exists a partition α of X (modulo 0) by sets of positive measure, such that

(1) Markov: for all $a \in \alpha$, $T(a)$ is a union (modulo 0) of elements of α and $T : a \to T(a)$ is invertible.
(2) Big image and preimage property: there exists a subset $\{a_1, \dots, a_n\}$ of α with the following property: for any $a \in \alpha$, there exist $i, j \in \{1, \dots, n\}$ such that $a \subset T(a_i)$ and $a_j \subset T(a)$ (modulo 0).
(3) Expansion: there exists $\gamma < 1$ such that for all $a \in \alpha$, for almost all $x, y \in a$, $d(Tx, Ty) \geq \gamma^{-1} d(x, y)$.
(4) Distortion: for $a \in \alpha$, let g be the inverse of the jacobian of T on a, i.e.,

$$g(x) = \frac{\mathrm{d}m_{|a}}{\mathrm{d}(m \circ T_{|a})}(x) \quad \text{for} \quad x \in a.$$

Then there exists C such that, for all $a \in \alpha$, for almost all $x, y \in a$,

$$\left|1 - \frac{g(x)}{g(y)}\right| \leq Cd(Tx, Ty).$$

A Gibbs–Markov map is **mixing** if, for all $a, b \in \alpha$, there exists N such that $b \subset T^n(a) \mod 0$ for any $n > N$. Since the general case reduces to the mixing one, we will only consider mixing Gibbs–Markov maps.

For $f : X \to \mathbb{R}$ and $A \subset X$, let $Df(A)$ denote the best Lipschitz constant of f on A. If f is integrable, we will write $\int f$ or $E(f)$ for $\int f \, \mathrm{d}m$, the reference measure being always $\mathrm{d}m$. Our main result follows.

Theorem 1.5: *Let $T : X \to X$ be a mixing probability preserving a Gibbs–Markov map, and let $f : X \to \mathbb{R}$ satisfy $\sum_{a\in\alpha} m(a)Df(a)^\eta < \infty$ for some $\eta \in (0, 1]$.*

Assume $f \in L^2$. Then

- *Either f is the sum of a measurable coboundary and a constant, i.e., there exist a measurable function u and a real number c such that $f = u - u \circ T + c$ almost everywhere. Then u is bounded, and $S_n f - nc$ converges in distribution to the difference $Z - Z'$ where Z and Z' are independent random variables with the same distribution as u.*
- *Otherwise, let $\tilde{f} = f - \int f \, \mathrm{d}m$, and define $\sigma^2 = \int \tilde{f}^2 + 2\sum_{k=1}^{\infty} \int \tilde{f} \cdot \tilde{f} \circ T^k$. Then this series converges, $\sigma^2 > 0$, and $(S_n f - n \int f)/\sqrt{n}$ converges in distribution to $\mathcal{N}(0, \sigma^2)$.*

Assume that f does not belong to L^2. Let $Z_0, Z_1, \ldots$ be i.i.d. random variables with the same distribution as f. Consider sequences $A_n \in \mathbb{R}$ and $B_n > 0$, and a nondegenerate random variable W. Then $(S_n f - A_n)/B_n$ converges to W if and only if $(\sum_{k=0}^{n-1} Z_k - A_n)/B_n$ converges to W.

In particular, it follows from the classification in the i.i.d. case that the Birkhoff sums of a function f satisfy a nondegenerate limit theorem if and only if the distribution of f belongs to the class $\mathcal{D}$ described in Paragraph 1.1.

In the L^2 case, the behavior of Birkhoff sums can be quite different from the i.i.d. case (see the formula for σ^2, encompassing the interactions between different times). On the other hand, when $f \notin L^2$, the behavior is exactly the same as in the i.i.d. case (the interactions are negligible with respect to the growth of the sums), and the good scaling coefficients can be read directly from the independent case Theorem 1.3.

The "sufficiency" part of the theorem (i.e., the convergence of the Birkhoff sums if f is in the domain of attraction of a gaussian or stable law) is known under stronger regularity assumptions: if the function f is locally Hölder continuous (i.e. $\sup_{a\in\alpha} Df(a) < \infty$), then the result is proved in [AD01b, AD01a] for $f \notin L^2$, and it follows from the classical Nagaev method (see, e.g., [RE83, GH88] for subshifts of finite type) when $f \in L^2$. The article [Gou04] proves the same results under the slightly weaker assumption $\sum m(a)Df(a) < \infty$. However, these methods are not sufficient to deal with the weaker assumption $\sum m(a)Df(a)^\eta < \infty$, hence new arguments will be required to prove the sufficiency part of Theorem 1.5. The main difficulty is the following: even if f belongs to all L^p spaces and $\sum m(a)Df(a)^\eta < \infty$, it is possible that $\hat{T}f$ is not locally Hölder continuous, in the sense that there exists $a \in \alpha$ with $D(\hat{T}f)(a) = \infty$ (here, $\hat{T}$ denotes the transfer operator associated to T).[1]

However, the main novelty of the previous theorem is the necessity part, showing that no exotic limit theorem can hold for Gibbs–Markov maps, even if one assumes only very weak regularity of the observable.

Remark 1.6: The regularity condition $\sum_{a\in\alpha} m(a)Df(a)^\eta < \infty$ is weaker than the conditions usually encountered in the literature, but it appears in some natural examples: for instance, if one tries to prove limit theorems for the observable $f_0(x) = x^{-a}$ under the iteration of $x \mapsto 2x \mod 1$, by inducing on $[1/2, 1]$, then the resulting induced observable f satisfies such a condition for some $\eta = \eta(a) < 1$, but not for $\eta = 1$.

Remark 1.7: For general transitive Gibbs–Markov maps (without the mixing assumption), it is still possible to prove that, if the Birkhoff sums $S_n f$ of a function f (with $\sum m(a)Df(a)^\eta < \infty$) satisfy a nondegenerate limit theorem, then the distribution of f belongs to the class $\mathcal{D}$: the proof we shall give below also applies to merely transitive maps. However, the converse is not true. More precisely, functions in $\mathcal{D}$ which are not the sum of a coboundary and a constant satisfy a limit theorem, just like in the mixing case, but this is not the case in general for coboundaries.

[1] This is, for instance, the case if T is the full Markov shift on infinitely many symbols $a_0, a_1, \dots$ with $m(a_i) = Ce^{-i^2/2}$, and f vanishes on $[a_i]$ but on $[a_i] \cap \bigcap_{n=1}^{i-1} T^{-n}(a_0)$, where it is equal to e^{i^2}.

1.3. A more general setting. Our results on Gibbs–Markov maps will be a consequence of a more general theorem, making it possible to obtain necessary and sufficient conditions for limit theorems of Birkhoff sums whenever one can obtain sufficiently precise information on characteristic functions.

Definition 1.8: Let $T : X \to X$ be a probability preserving mixing map, and let $f : X \to \mathbb{R}$ be measurable. The function f admits a **characteristic expansion** if there exist a neighborhood I of 0 in $\mathbb{R}$, two measurable functions $\lambda, \mu : I \to \mathbb{R}$ continuous at 0 with $\lambda(0) = \mu(0) = 1$, and a sequence ϵ_n tending to 0 such that, for any $t \in I$ and any $n \in \mathbb{N}$,

$$\left|E(e^{\mathbf{i}tS_nf}) - \lambda(t)^n\mu(t)\right| \leq \epsilon_n. \tag{1.3}$$

This characteristic expansion is **accurate** if one of the following properties holds:

- Either there exist $q \leq 2$ and $\epsilon > 0$ such that $f \notin L^q$ and

$$\lambda(t) = E(e^{\mathbf{i}tf}) + O(|t|^{q+\epsilon}) + O(t^2) + o\left(\int |e^{\mathbf{i}tf} - 1|^2\right) + O\left(\int |e^{\mathbf{i}tf} - 1|\right)^2. \tag{1.4}$$

- Or $f \in L^2$ and there exists $c \in \mathbb{C}$ such that

$$\lambda(t) = 1 + \mathbf{i}tE(f) - ct^2/2 + o(t^2). \tag{1.5}$$

When $f \notin L^2$, this definition tells that a characteristic expansion is accurate if $\lambda(t)$ is close to $E(e^{\mathbf{i}tf})$, up to error terms described by (1.4). It should be noted that these error terms are *not* always negligible with respect to $1 - E(e^{\mathbf{i}tf})$, but they are nevertheless sufficiently small (for sufficiently many values of t) to ensure a good behavior, as shown by the following theorem.

Theorem 1.9: *Let $T : X \to X$ be a probability preserving mixing map, and let $f : X \to \mathbb{R}$ admit an accurate characteristic expansion.*

Assume that $f \in L^2$. Let $\lambda(t) = 1+\mathbf{i}tE(f)-ct^2/2+o(t^2)$ be the characteristic expansion of f. Then $\sigma^2 := c - E(f)^2 \geq 0$, and $(S_nf - nE(f))/\sqrt{n}$ converges in distribution to $\mathcal{N}(0, \sigma^2)$.

Assume that $f \notin L^2$. Let $Z_0, Z_1, \dots$ be i.i.d. random variables with the same distribution as f. Consider sequences $A_n \in \mathbb{R}$ and $B_n > 0$, and a nondegenerate random variable W. Then $(S_nf - A_n)/B_n$ converges to W if and only if $(\sum_{k=0}^{n-1} Z_k - A_n)/B_n$ converges to W.

The flavor of this theorem is very similar to Theorem 1.5. The only difference is in the L^2 case, when $\sigma^2 = 0$: Theorem 1.9 only says that $(S_n f - nE(f))/\sqrt{n}$ converges in distribution to 0 (note that this is a *degenerate* limit theorem) while Theorem 1.5 gives a more precise conclusion in this case, showing that $S_n f - nE(f)$ converges in distribution to a nontrivial random variable. To get this conclusion, one needs to show that a function f satisfying $\sigma^2 = 0$ is a coboundary — this is indeed the case for Gibbs–Markov maps, as we will see in Paragraph 3.6.

To deduce Theorem 1.5 from Theorem 1.9, we should of course check the assumptions of the latter theorem. The following proposition is therefore the core of our argument concerning Gibbs–Markov maps.

PROPOSITION 1.10: *Let T be a mixing Gibbs–Markov map, and let $f : X \to \mathbb{R}$ satisfy $\sum_{a\in\alpha} m(a)Df(a)^\eta < \infty$ for some $\eta > 0$. Then f admits an accurate characteristic expansion.*

Remark 1.11: If we strengthened Definition 1.8, by requiring for instance $\lambda(t) = E(e^{\mathbf{i}tf}) + O(t^2)$ when $f \not\in L^2$, then Theorem 1.9 would be much easier to prove. However, we would not be able to prove Proposition 1.10 with this stronger definition. The form of the error term in (1.4) is the result of a tradeoff between what is sufficient to prove Theorem 1.9, and what we can prove for Gibbs–Markov maps.

The rest of the paper is organized as follows: Section 2 is devoted to the proof of Theorem 1.9, using general considerations on characteristic functions, while the results concerning Gibbs–Markov maps (Proposition 1.10 and Theorem 1.5) are proved in Section 3. The required characteristic expansion is obtained in some cases using classical perturbation theory as in [AD01b], but other tools are also required in other cases: weak perturbation theory [KL99, GL06, HP08] and interpolation spaces [BL76]. Finally, Appendix A describes another application of our techniques, to the speed in the central limit theorem.

2. Using accurate characteristic expansions

In this section, we prove Theorem 1.9. Let f be a function satisfying an accurate characteristic expansion.

Assume first that f is square integrable. Let $t \in \mathbb{R}$. If n is large enough, $t/\sqrt{n}$ belongs to the domain of definition of λ, and

$$\lambda\left(\frac{t}{\sqrt{n}}\right)^n = \exp\Big(n\log\big[1+\mathbf{i}tE(f)/\sqrt{n} - ct^2/(2n) + o(1/n)\big]\Big)$$
$$= \exp\Big(n\big[\mathbf{i}tE(f)/\sqrt{n} - ct^2/(2n) + t^2E(f)^2/(2n) + o(1/n)\big]\Big).$$

Hence, $e^{-\mathbf{i}t\sqrt{n}E(f)}\lambda(t/\sqrt{n})^n$ converges to $e^{-(c-E(f)^2)t^2/2}$. By definition of a characteristic expansion, this implies that $e^{-\mathbf{i}t\sqrt{n}E(f)}E(e^{\mathbf{i}tS_nf/\sqrt{n}})$ converges to $e^{-(c-E(f)^2)t^2/2}$. Therefore, $e^{-(c-E(f)^2)t^2/2}$ is the characteristic function of a random variable W and $(S_nf - nE(f))/\sqrt{n}$ converges in distribution to W. This yields $\sigma^2 := c - E(f)^2 \geq 0$, and $W = \mathcal{N}(0,\sigma^2)$, as desired. The proof of Theorem 1.9 is complete in this case.

We now turn to the other more interesting case, where $f \notin L^2$. We have apparently two different implications to prove, but we will prove them at the same time, using the following proposition.

Proposition 2.1: *Let $T : X \to X$ and $\tilde{T} : \tilde{X} \to \tilde{X}$ be two probability preserving mixing maps, and let $f : X \to \mathbb{R}$ and $\tilde{f} : \tilde{X} \to \mathbb{R}$ be two functions with the same distribution. Assume that both of them admit an accurate characteristic expansion, and do not belong to L^2. If $(\sum_{k=0}^{n-1} f\circ T^k - A_n)/B_n$ converges in distribution to a nondegenerate random variable W, then $(\sum_{k=0}^{n-1}\tilde{f}\circ\tilde{T}^k - A_n)/B_n$ also converges to W.*

Let us show how this proposition implies Theorem 1.9.

Conclusion of the proof of Theorem 1.9, assuming Proposition 2.1. Let $f \notin L^2$ admit an accurate characteristic expansion.

Let $\tilde{T}$ be the left shift on the space $\tilde{X} = \mathbb{R}^{\mathbb{N}}$, and let $\tilde{f}(x_0, x_1, \dots) = x_0$. We endow $\tilde{X}$ with the product measure such that $\tilde{f}, \tilde{f}\circ\tilde{T}, \dots$ are i.i.d. and distributed as f. Then $\tilde{f}$ admits an accurate characteristic expansion (with $\tilde{\mu}(t) = 1$ and $\tilde{\lambda}(t) = E(e^{\mathbf{i}tf})$).

Proposition 2.1 shows that the convergence of $(S_nf - A_n)/B_n$ to W gives the convergence of $(\sum_{k=0}^{n-1}\tilde{f}\circ\tilde{T}^k - A_n)/B_n$ to W. This is one of the desired implications in Theorem 1.9. The other implication follows from the same argument, but exchanging the roles of T and $\tilde{T}$ in Proposition 2.1. ∎

The rest of this section is devoted to the proof of Proposition 2.1. We fix once and for all $T, \tilde{T}$ and $f, \tilde{f}$ as in the assumptions of this proposition, and assume that $(S_n f - A_n)/B_n$ converges in distribution to a nondegenerate random variable W. Let us also fix q and ϵ such that $f \notin L^q$ and $\lambda(t), \tilde{\lambda}(t)$ satisfy (1.4) (if the values of q and ϵ do not coincide for the expansions of $\lambda(t)$ and $\tilde{\lambda}(t)$, just take the minimum of the two).

Let $\Phi(t) = E(1 - \cos(tf)) \geq 0$; this function will play an essential role in the following arguments.

LEMMA 2.2: *We have $\int |e^{\mathbf{i}tf} - 1|^2 = 2\Phi(t)$. Moreover, since f does not belong to L^2,*

$$t^2 + \left(\int |e^{\mathbf{i}tf} - 1|\right)^2 = o(\Phi(t)) \quad \text{when } t \to 0. \tag{2.1}$$

Proof. Writing $|e^{\mathbf{i}tf} - 1|^2 = (e^{\mathbf{i}tf} - 1)(e^{-\mathbf{i}tf} - 1)$ and expanding the product, the first assertion of the lemma is trivial.

To prove that $t^2 = o(\int |e^{\mathbf{i}tf} - 1|^2)$, let us show that

$$\int \left|\frac{e^{\mathbf{i}tf} - 1}{t}\right|^2 \mathrm{d}m \to +\infty \quad \text{when } t \to 0. \tag{2.2}$$

The integrand converges to $|f|^2$, whose integral is infinite. Since a sequence f_n of nonnegative functions always satisfies $\int \liminf f_n \leq \liminf \int f_n$, by Fatou's Lemma, we get (2.2).

Let us now check that $\left(\int |e^{\mathbf{i}tf} - 1|\right)^2 = o(\int |e^{\mathbf{i}tf} - 1|^2)$. Fix a large number M and partition the space into $A_M = \{|f| \leq M\}$ and $B_M = \{|f| > M\}$. Since $(a+b)^2 \leq 2a^2 + 2b^2$ for any $a, b \geq 0$, we get

$$\begin{aligned}
\left(\int |e^{\mathbf{i}tf} - 1|\right)^2 &= \left(\int_{A_M} |e^{\mathbf{i}tf} - 1| + \int_{B_M} |e^{\mathbf{i}tf} - 1|\right)^2 \\
&\leq 2\left(\int_{A_M} |e^{\mathbf{i}tf} - 1|\right)^2 + 2\left(\int_{B_M} |e^{\mathbf{i}tf} - 1|\right)^2 \\
&\leq 2M^2t^2 + 2\left\|1_{B_M}\right\|_{L^2}^2 \left\|e^{\mathbf{i}tf} - 1\right\|_{L^2}^2,
\end{aligned}$$

by Cauchy–Schwarz inequality. The term $2M^2t^2$ is negligible with respect to $\int |e^{\mathbf{i}tf} - 1|^2$, by (2.2), while the second term is $m(B_M) \int |e^{\mathbf{i}tf} - 1|^2$. Choosing M large enough, we can ensure that $m(B_M)$ is arbitrarily small, concluding the proof. ∎

Lemma 2.2 shows that (1.4) is equivalent to

$$\lambda(t) = E(e^{\mathbf{i}tf}) + O(|t|^{q+\epsilon}) + o(\Phi(t)). \tag{2.3}$$

The main difficulty is that, for a general function f not belonging to L^q, $|t|^{q+\epsilon}$ is not always negligible with respect to $\Phi(t)$. This is, however, true along a subsequence of t's:

Lemma 2.3: *Since f does not belong to L^q, there exists an infinite set $A \subset \mathbb{N}$ such that, for any $t \in \Lambda := \bigcup_{n\in A}[2^{-n-1}, 2^{-n}]$,*

$$|t|^{q+\epsilon/2} \leq \Phi(t). \tag{2.4}$$

Proof. Assume by contradiction that, for any large enough n, there exists $t_n \in [2^{-n-1}, 2^{-n}]$ with $\int_X 1 - \cos(t_n f) < |t_n|^{q+\epsilon/2}$. If $x \in X$ is such that $|f(x)| \in [2^{n-1}, 2^n]$, then $|t_n f(x)| \in [1/4, 1]$. Since $1 - \cos(y)$ is bounded from below by $c > 0$ on $[-1, -1/4] \cup [1/4, 1]$, we get

$$m\{|f| \in [2^{n-1}, 2^n]\} \leq c^{-1} \int 1 - \cos(t_n f) \leq C|t_n|^{q+\epsilon/2} \leq C2^{-n(q+\epsilon/2)}.$$

Hence, $\sum 2^{qn} m\{|f| \in [2^{n-1}, 2^n]\}$ is finite. This implies that f belongs to L^q, a contradiction. ∎

Lemma 2.4: *Along Λ, we have $|\lambda(t)|^2 = 1 - (2 + o(1))\Phi(t)$.*

Proof. Along Λ, the previous lemma and (2.3) give $\lambda(t) = E(e^{\mathbf{i}tf}) + o(\Phi(t))$. Hence,

$$\begin{aligned} |\lambda(t)|^2 &= |E(e^{\mathbf{i}tf})|^2 + o(\Phi(t)) \\ &= (1 - E(1 - e^{\mathbf{i}tf})) \cdot (1 - E(1 - e^{-\mathbf{i}tf})) + o(\Phi(t)) \\ &= 1 - 2E(1 - \cos(tf)) + |E(1 - e^{\mathbf{i}tf})|^2 + o(\Phi(t)). \end{aligned}$$

Moreover, $E(1 - \cos(tf)) = \Phi(t)$, and $|E(1 - e^{\mathbf{i}tf})|^2 \leq \left(\int |1 - e^{\mathbf{i}tf}|\right)^2$, which is negligible with respect to $\Phi(t)$ by Lemma 2.2. This proves the lemma. ∎

Lemma 2.5: *The sequence B_n tends to infinity.*

Proof. Assume by contradiction that B_n does not tend to infinity. Then, there exists a subsequence $j(n)$ such that the distribution of $S_{j(n)}f - A_{j(n)}$ is tight. Since T is mixing, [AW00, Theorem 2] implies the existence of $c \in \mathbb{R}$ and of a measurable function $u : X \to \mathbb{R}$ such that $f = u - u \circ T + c$ almost everywhere. In particular, $S_n f - nc$ converges in distribution, to $Z := Z_1 - Z_2$ where Z_1 and

Z_2 are i.i.d. and distributed as u. Hence, $e^{-\mathbf{i}tnc}E(e^{\mathbf{i}tS_nf})$ converges to $E(e^{\mathbf{i}tZ})$, and therefore $|E(e^{\mathbf{i}tS_nf})| \to |E(e^{\mathbf{i}tZ})|$. However, $E(e^{\mathbf{i}tS_nf}) = \mu(t)\lambda(t)^n + o(1)$. If $|\lambda(t)| < 1$, we obtain $E(e^{\mathbf{i}tZ}) = 0$.

Along Λ, the function Φ is positive (by (2.4)) and $|\lambda(t)|^2 = 1-(2+o(1))\Phi(t)$ by the previous lemma. Hence, if t is small enough and belongs to Λ, we have $|\lambda(t)| < 1$, and $E(e^{\mathbf{i}tZ}) = 0$. In particular, the function $t \mapsto E(e^{\mathbf{i}tZ})$ is not continuous at 0, which is a contradiction since a characteristic function is always continuous. ∎

Lemma 2.6: *The sequence B_{n+1}/B_n converges to* 1.

Proof. We know that $(S_nf - A_n)/B_n$ converges in distribution to a nondegenerate random variable W. Since the measure is invariant, $(S_nf\circ T - A_n)/B_n$ also converges to W. Since $B_n \to \infty$, this implies that $(S_{n+1}f - A_n)/B_n$ converges to W. However, $(S_{n+1}f - A_{n+1})/B_{n+1}$ converges to W. The convergence of types theorem (Theorem 1.4) therefore yields $B_{n+1}/B_n \to 1$. ∎

Slowly varying functions have been defined in Definition 1.2.

Lemma 2.7: *There exist $d \in (0,2]$ and a slowly varying function L such that $B_n \sim n^{1/d}L(n)$.*

Proof. Since $B_n \to \infty$, the convergence $(S_nf - A_n)/B_n \to W$ translates into: $e^{-\mathbf{i}tA_n/B_n}\lambda(t/B_n)^n \to E(e^{\mathbf{i}tW})$ uniformly on small neighborhoods of 0. Hence, $|\lambda(t/B_n)|^{2n} \to |E(e^{\mathbf{i}tW})|^2 = E(e^{\mathbf{i}tZ})$ where $Z := W - W'$ is the difference of two independent copies of W. Taking the logarithm, we get

$$2n\log|\lambda(t/B_n)| \to \log E(e^{\mathbf{i}tZ}). \tag{2.5}$$

Since $B_n \to \infty$ and $B_{n+1}/B_n \to 1$, [BGT87, Proposition 1.9.4] implies that, for $t > 0$, one can write $|\lambda(t)|^2 = \exp(-t^dL_0(1/t))$ for some slowly varying function L_0 and some real number d. Moreover, $E(e^{\mathbf{i}tZ}) = e^{-ct^d}$ for some $c > 0$. Since $E(e^{\mathbf{i}tZ})$ is a characteristic function, this restricts the possible values of d to $d \in (0,2]$.

Let $t_0 > 0$ be such that $E(e^{\mathbf{i}t_0Z}) \in (0,1)$. The convergence (2.5) for $t = t_0$ becomes $n \sim CB_n^d/L_0(B_n)$ for some $C > 0$. Since $d > 0$, the function $x \mapsto Cx^d/L_0(x)$ is asymptotically invertible by [BGT87, Theorem 1.5.12], and admits an inverse of the form $x \mapsto x^{1/d}L(x)$ where L is slowly varying. We get $B_n \sim n^{1/d}L(n)$. ∎

Lemma 2.8: *The number d given by Lemma 2.7 satisfies $d \leq q + \epsilon/2$.*

Proof. Let $t_0 > 0$ satisfy $E(e^{\mathbf{i}t_0 Z}) \in (0,1)$. The sequence $|\lambda(t_0/B_n)|^{2n}$ converges to $E(e^{\mathbf{i}t_0 Z})$. Taking logarithms, we obtain the existence of $a > 0$ such that

$$-n \log \left|\lambda\left(\frac{t_0}{B_n}\right)\right|^2 \to a. \tag{2.6}$$

Since $B_{n+1}/B_n \to 1$, there exists $j(n) \to \infty$ such that $t_0/B_{j(n)} \in \Lambda$ (where Λ is defined in Lemma 2.3). Along this sequence, we have $|\lambda(t_0/B_{j(n)})|^2 = 1 - (2 + o(1))\Phi(t_0/B_{j(n)})$ by Lemma 2.4. Taking the logarithm and using (2.6), we obtain $j(n)\Phi(t_0/B_{j(n)}) \to a/2$. By (2.4), this yields $j(n)/B_{j(n)}^{q+\epsilon/2} = O(1)$. Moreover, by Lemma 2.7,

$$j(n)/B_{j(n)}^{q+\epsilon/2} \sim j(n)^{1-(q+\epsilon/2)/d}/L(j(n))^{q+\epsilon/2}. \tag{2.7}$$

This sequence can be bounded only if $1 - (q + \epsilon/2)/d \leq 0$, concluding the proof. ∎

Proof of Proposition 2.1. For small enough t, $|\lambda(t) - 1| < 1/2$. Hence, it is possible to define $\log \lambda(t)$ by the series $\log(1 - s) = -\sum s^k/k$. Since the logarithm is a Lipschitz function, (2.3) gives

$$\log \lambda(t) = \log E(e^{\mathbf{i}tf}) + O(|t|^{q+\epsilon}) + o(\Phi(t)).$$

Moreover, $1 - E(e^{\mathbf{i}tf}) = \Phi(t) - \mathbf{i}E(\sin tf)$, hence

$$-\log E(e^{\mathbf{i}tf}) = \Phi(t) - \mathbf{i}E(\sin tf) + o(\Phi(t)) + O(|E(\sin tf)|^2). \tag{2.8}$$

Moreover, $|E(\sin(tf))| = |\operatorname{Im} E(e^{\mathbf{i}tf} - 1)| \leq E|e^{\mathbf{i}tf} - 1|$. Using (2.1), we obtain $|E(\sin(tf))|^2 = o(\Phi(t))$. We have proved

$$-\log \lambda(t) = \Phi(t) - \mathbf{i}E(\sin tf) + o(\Phi(t)) + O(|t|^{q+\epsilon}). \tag{2.9}$$

The convergence $(S_n f - A_n)/B_n \to W$ also reads

$$e^{-\mathbf{i}tA_n/B_n}\lambda(t/B_n)^n \to E(e^{\mathbf{i}tW}).$$

By (2.9), the left hand side is

$$\exp\Big(-\mathbf{i}tA_n/B_n - n\Phi(t/B_n) + n\mathbf{i}E(\sin(tf/B_n)) + o(n\Phi(t/B_n)) + O(n/B_n^{q+\epsilon})\Big).$$

By Lemma 2.8, $n/B_n^{q+\epsilon}$ tends to 0 when $n \to \infty$. Hence, the last equation can also be written as

$$\exp\Big(-\mathbf{i}tA_n/B_n - n\Phi(t/B_n) + n\mathbf{i}E(\sin(tf/B_n)) + o(n\Phi(t/B_n)) + o(1)\Big). \tag{2.10}$$

To prove the desired convergence of $(\tilde{S}_n\tilde{f} - A_n)/B_n$ to W, we should prove that $e^{-\mathbf{i}tA_n/B_n}\tilde{\lambda}(t/B_n)^n$ converges to $E(e^{\mathbf{i}tW})$. The previous arguments also apply to $\tilde{\lambda}$, and show that

$$(2.11)\quad e^{-\mathbf{i}tA_n/B_n}\tilde{\lambda}\left(\frac{t}{B_n}\right)^n = \exp\Big(-\mathbf{i}tA_n/B_n - n\Phi(t/B_n) + n\mathbf{i}E(\sin(tf/B_n)) + \tilde{o}(n\Phi(t/B_n)) + \tilde{o}(1)\Big),$$

where we have used the notation $\tilde{o}$ to emphasize the fact that these negligible terms may be different from those in (2.10).

Let us now conclude the proof by showing that (2.11) converges to $E(e^{\mathbf{i}tW})$, using the fact that (2.10) converges to $E(e^{\mathbf{i}tW})$. The only possible problem comes from the negligible term $\tilde{o}(n\Phi(t/B_n))$ (since the term $\tilde{o}(1)$ has no influence on the limit).

We treat two cases. Assume first that $E(e^{\mathbf{i}tW}) \neq 0$. Then the modulus of $\lambda(t/B_n)^n$ converges to a nonzero real number. In particular, $n\Phi(t/B_n)$ converges, which implies that $\tilde{o}(n\Phi(t/B_n))$ converges to 0. This concludes the proof in this case.

Assume now that $E(e^{\mathbf{i}tW}) = 0$. This implies that the modulus of $\lambda(t/B_n)^n$ converges to 0. By (2.10), this yields $n\Phi(t/B_n) \to +\infty$. In this case, we have no control on the argument of $e^{-\mathbf{i}tA_n/B_n}\tilde{\lambda}(t/B_n)^n$ (since the term $\tilde{o}(n\Phi(t/B_n))$ may very well not tend to 0), but its modulus tends to 0. This is sufficient to get again $e^{-\mathbf{i}tA_n/B_n}\tilde{\lambda}(t/B_n)^n \to 0 = E(e^{\mathbf{i}tW})$. This concludes the proof of Proposition 2.1. ∎

3. Characteristic expansions for Gibbs–Markov maps

3.1. The accurate characteristic expansion for non-integrable functions. Let us fix a mixing probability preserving Gibbs–Markov map $T: X \to X$, as well as a measurable function $f: X \to \mathbb{R}$ with $\sum m(a)Df(a)^\eta < \infty$ for some $\eta \in (0,1]$.

Let $\hat{T}$ denote the transfer operator associated to T (defined by duality by $\int u \cdot v \circ T \,\mathrm{d}m = \int \hat{T}u \cdot v \,\mathrm{d}m$). It is given explicitly by

$$(3.1)\qquad \hat{T}u(x) = \sum_{Ty=x} g(y)u(y),$$

where g is the inverse of the jacobian of T. We will need the following inequality: there exists a constant C such that

$$C^{-1}m(a) \leq g(x) \leq Cm(a) \tag{3.2}$$

for any $a \in \alpha$ and $x \in a$. This follows from the assumption of bounded distortion for Gibbs–Markov maps.

Let $\mathcal{L}$ be the space of bounded functions $u : X \to \mathbb{C}$ such that

$$\sup_{a\in\alpha} \sup_{x,y\in a} |u(x) - u(y)|/d(x, y)^\eta < \infty. \tag{3.3}$$

Then $\hat{T}$ acts continuously on $\mathcal{L}$, has a simple eigenvalue at 1 and the rest of its spectrum is contained in a disk of radius < 1. Moreover, it satisfies an inequality

$$\left\|\hat{T}^n u\right\|_{\mathcal{L}} \leq C\gamma^n \|u\|_{\mathcal{L}} + C \|u\|_{L^1}, \tag{3.4}$$

for some $\gamma < 1$. This follows from [AD01b, Proposition 1.4 and Theorem 1.6].

Let us now define a perturbed transfer operator $\hat{T}_t$ by $\hat{T}_t(u) = \hat{T}(e^{\mathbf{i}tf}u)$. Using the estimate $\sum m(a)Df(a)^\eta < \infty$, one can check that the operator $\hat{T}_t$ acts continuously on $\mathcal{L}$, and

$$\left\|\hat{T}_t - \hat{T}\right\|_{\mathcal{L}\to\mathcal{L}} = O(|t|^\eta + E|e^{\mathbf{i}tf} - 1|). \tag{3.5}$$

This follows from Lemma 3.5 and the proof of Corollary 3.6 in [Gou04].

The estimate (3.5) is a strong continuity estimate. We can therefore apply the following classical perturbation theorem (which follows, for instance, from [Kat66, Sections III.6.4 and IV.3.3]).

Theorem 3.1: *Let A be a continuous operator on a Banach space $\mathcal{B}$, for which 1 is a simple eigenvalue, and the rest of its spectrum is contained in a disk of radius < 1. Let A_t (for small enough t) be a family of continuous operators on $\mathcal{B}$, such that $\|A_t - A\|_{\mathcal{B}\to\mathcal{B}} \to 0$ when $t \to 0$.*

Then, for any small enough t, there exists a decomposition $E_t \oplus F_t$ of $\mathcal{B}$ into a one-dimensional subspace and a closed hyperplane, such that E_t and F_t are invariant under A_t. Moreover, A_t is the multiplication by a scalar $\lambda(t)$ on E_t, while $\|(A_t)^n_{|F_t}\|_{\mathcal{B}\to\mathcal{B}} \leq C\gamma^n$ for some $\gamma < 1$ and $C > 0$.

The eigenvalue $\lambda(t)$ and the projection P_t on E_t with kernel F_t satisfy

$$|\lambda(t) - 1| \leq C \|A_t - A\|_{\mathcal{B}\to\mathcal{B}} \tag{3.6}$$

and

$$\|P_t - P_0\|_{\mathcal{B}\to\mathcal{B}} \leq C\,\|A_t - A\|_{\mathcal{B}\to\mathcal{B}}\,. \tag{3.7}$$

This theorem yields an eigenvalue $\lambda(t)$ of $\hat{T}_t$ for small t, and an eigenfunction $\xi_t = P_t 1/\int P_t 1$ such that $\int \xi_t = 1$ and

$$\|\xi_t - 1\|_{\mathcal{L}} = O(|t|^\eta + E|e^{\mathbf{i}tf} - 1|), \tag{3.8}$$

by (3.7).

We have

$$E(e^{\mathbf{i}tS_n f}) = \int \hat{T}_t^n(1) = \lambda(t)^n \int P_t 1 + O(\gamma^n) = \mu(t)\lambda(t)^n + O(\gamma^n), \tag{3.9}$$

for $\mu(t) = \int P_t 1$. This proves that f admits a characteristic expansion. To prove Proposition 1.10, we have to show that this expansion is accurate, i.e., to get precise estimates on $\lambda(t)$. We have

$$\lambda(t) = \int \lambda(t)\xi_t = \int \hat{T}_t \xi_t = \int \hat{T}_t 1 + \int (\hat{T}_t - \hat{T})(\xi_t - 1), \tag{3.10}$$

hence

$$\lambda(t) = E(e^{\mathbf{i}tf}) + \int (e^{\mathbf{i}tf} - 1)(\xi_t - 1). \tag{3.11}$$

When $\eta = 1$ (i.e., $\sum m(a)Df(a) < \infty$) and $f \notin L^2$, (3.11) together with the estimate (3.8) readily imply that the characteristic expansion of f is accurate, concluding the proof of Proposition 1.10 in this case. The general case requires more work.

We first deal with the case $f \notin L^{1+\eta/2}$. In this case, we already have enough information to conclude:

Lemma 3.2: *If $f \notin L^{1+\eta/2}$, then f admits an accurate characteristic expansion.*

Proof. The equation (3.11) together with (3.8) yield

$$\lambda(t) = E(e^{\mathbf{i}tf}) + O\left(|t|^\eta \cdot \int |e^{\mathbf{i}tf} - 1|\right) + O\left(\int |e^{\mathbf{i}tf} - 1|\right)^2. \tag{3.12}$$

Let $p \in [0,1]$ be such that $f \in L^p$ and $f \notin L^{p+\eta/2}$ (we use the convention that every measurable function belongs to L^0). For any $x \in \mathbb{R}$, $|e^{\mathbf{i}x} - 1| \leq 2|x|^p$. Then

$$|t|^\eta \int |e^{\mathbf{i}tf} - 1| \leq 2|t|^\eta \int |t|^p |f|^p \leq C|t|^{p+\eta}. \tag{3.13}$$

This yields the accurate characteristic expansion (1.4) as desired, for $q = p+\eta/2$ and $\epsilon = \eta/2$. ∎

The case where $f \in L^{1+\eta/2}$ is a lot trickier. It requires a more general spectral perturbation theorem, essentially due to Keller and Liverani. Unfortunately, this theorem is sufficient only when there exists $q < 2$ such that $f \notin L^q$, while the remaining case can only be treated using a generalization of this theorem, involving several successive derivatives of the operators, that we will describe in the next paragraph.

3.2. A GENERAL SPECTRAL THEOREM. In this paragraph, we describe a general spectral theorem extending the results of [KL99] to the case of several derivatives. A very similar result has been proved in [GL06], but with slightly stronger assumptions that will not be satisfied in the forthcoming application to Gibbs–Markov maps (in particular, [GL06] requires (3.16) below to hold for $0 \le i < j \le N$, instead of $1 \le i < j \le N$). Let us also mention [HP08] for related results.

Let $\mathcal{B}_0 \supset \mathcal{B}_1 \supset \cdots \supset \mathcal{B}_N$, $N \in \mathbb{N}^*$, be a finite family of Banach spaces, let $I \subset \mathbb{R}$ be a fixed open interval containing 0, and let $\{A_t\}_{t\in I}$ be a family of operators acting on each of the above Banach spaces. Let also $b_0, b_1, \dots, b_{N-1} \in (0,1]$ (usually, $b_i = 1$ for $i \ge 1$). Let $b(i,j) = \sum_{k=i}^{j-1} b_k$ for $0 \le i \le j \le N$. Assume that

$$\exists M > 0, \forall t \in I, \quad \|A_t^n f\|_{\mathcal{B}_0} \le C M^n \|f\|_{\mathcal{B}_0} \tag{3.14}$$

and

$$\exists \gamma < M, \ \forall t \in I, \quad \|A_t^n f\|_{\mathcal{B}_1} \le C\gamma^n \|f\|_{\mathcal{B}_1} + CM^n \|f\|_{\mathcal{B}_0} . \tag{3.15}$$

Assume also that there exist operators $Q_1, \dots, Q_{N-1}$ satisfying the following properties:

$$\forall 1 \le i < j \le N, \quad \|Q_{j-i}\|_{\mathcal{B}_j \to \mathcal{B}_i} \le C \tag{3.16}$$

and, setting $\Delta_0(t) := A_t$ and $\Delta_j(t) := A_t - A_0 - \sum_{k=1}^{j-1} t^k Q_k$ for $j \ge 1$,

$$\forall t \in I, \ \forall 0 \le i \le j \le N, \quad \|\Delta_{j-i}(t)\|_{\mathcal{B}_j \to \mathcal{B}_i} \le C|t|^{b(i,j)}. \tag{3.17}$$

These assumptions mean that $t \mapsto A_t$ is continuous at $t = 0$ as a function from $\mathcal{B}_i$ to $\mathcal{B}_{i-1}$, and that $t \mapsto A_t$ even has a Taylor expansion of order $N-1$, but the differentials take their values in weaker spaces.

For $\varrho > \gamma$ and $\delta > 0$, denote by $V_{\delta,\varrho}$ the set of complex numbers z such that $|z| \geq \varrho$ and, for all $1 \leq k \leq N$, the distance from z to the spectrum of A_0 acting on $\mathcal{B}_k$ is $\geq \delta$.

THEOREM 3.3: *Given a family of operators* $\{A_t\}_{t\in I}$ *satisfying conditions* (3.14), (3.15), (3.16) *and* (3.17) *and setting*

$$R_N(t) := \sum_{k=0}^{N-1} t^k \sum_{\ell_1+\cdots+\ell_j=k} (z-A_0)^{-1}Q_{\ell_1}(z-A_0)^{-1}\dots(z-A_0)^{-1}Q_{\ell_j}(z-A_0)^{-1},$$

for all $z \in V_{\delta,\varrho}$ *and* t *small enough, we have*

$$\left\|(z-A_t)^{-1} - R_N(t)\right\|_{\mathcal{B}_N\to\mathcal{B}_0} \leq C|t|^{\kappa b_0 + b(1,N)}$$

where $\kappa = \log(\varrho/\gamma)/\log(M/\gamma)$.

Hence, the resolvent $(z-A_t)^{-1}$ depends on t in a $C^{\kappa b_0+b(1,N)}$ way at $t=0$, when viewed as an operator from $\mathcal{B}_N$ to $\mathcal{B}_0$.

Notice that one of the results of [KL99] in the present setting reads

$$\left\|(z-A_t)^{-1} - (z-A_0)^{-1}\right\|_{\mathcal{B}_1\to\mathcal{B}_0} \leq C|t|^{\kappa b_0}. \tag{3.18}$$

Accordingly, one has Theorem 3.3 in the case $N=1$ where no assumption is made on the existence of the operators Q_j.

We will use the following estimate of [KL99]:

LEMMA 3.4: *For any small enough* τ *and any* $z \in V_{\delta,\varrho}$, *we have*

$$\left\|(z-A_0)^{-1}u\right\|_{\mathcal{B}_0} \leq C\tau^{\kappa}\|u\|_{\mathcal{B}_1} + C\tau^{\kappa-1}\|u\|_{\mathcal{B}_0}. \tag{3.19}$$

Proof. This is essentially (11) in [KL99]. Let us recall the proof for the convenience of the reader. We have

$$(z-A_0)^{-1} = z^{-n}(z-A_0)^{-1}A_0^n + \frac{1}{z}\sum_{j=0}^{n-1}(z^{-1}A_0)^j \tag{3.20}$$

(this can be obtained for large enough z by taking the series expansion of $(z-A_0)^{-1}$ and isolating the first terms). Hence,

$$\begin{aligned}\left\|(z-A_0)^{-1}u\right\|_{\mathcal{B}_0} &\leq C|z|^{-n}\left\|(z-A_0)^{-1}\right\|_{\mathcal{B}_1\to\mathcal{B}_1}\left[\gamma^n\|u\|_{\mathcal{B}_1} + M^n\|u\|_{\mathcal{B}_0}\right] \\ &\quad + \frac{1}{|z|}\sum_{j=0}^{n-1}|z|^{-j}\left\|A_0^j\right\|_{\mathcal{B}_0\to\mathcal{B}_0}\|u\|_{\mathcal{B}_0} \\ &\leq C(\gamma/\varrho)^n\|u\|_{\mathcal{B}_1} + C(M/\varrho)^n\|u\|_{\mathcal{B}_0}.\end{aligned}$$

Let us choose n so that $(\gamma/\varrho)^n = \tau^\kappa$, i.e., $n = |\log\tau|/\log(M/\gamma)$. Then

$$(M/\varrho)^n = \exp\left(|\log\tau|\cdot\frac{\log(M/\varrho)}{\log(M/\gamma)}\right) = \tau^{\kappa-1}. \tag{3.21}$$ ∎

Proof of Theorem 3.3. We have

$$(z-A_t)^{-1} = (z-A_0)^{-1} + (z-A_t)^{-1}(A_t-A_0)(z-A_0)^{-1}. \tag{3.22}$$

If we want an expansion of $(z-A_t)^{-1}$ up to order $|t|^{\kappa b_0+b_1}$, this equation is sufficient: we can replace on the right $(z-A_t)^{-1}$ with $(z-A_0)^{-1}$ up to a small error $|t|^{\kappa b_0}$ (by (3.18)), and use the Taylor expansion of $A_t - A_0$ to conclude (since $A_t - A_0 = O_{\mathcal{B}_2\to\mathcal{B}_1}(|t|^{b_1})$, the global error is of order $|t|^{\kappa b_0+b_1}$). If we want a better precision $|t|^{\kappa b_0+b(1,N)}$, we should iterate the previous equation, so that in the end $(z-A_t)^{-1}$ is multiplied by a term of order $|t|^{b(1,N)}$.

This is done as follows. Let $A(z,t) := (A_t-A_0)(z-A_0)^{-1}$. Iterating the previous equation $N-1$ times, it follows that

(3.23)

$$\begin{aligned}(z-A_t)^{-1} &= \sum_{j=0}^{N-2}(z-A_0)^{-1}A(z,t)^j + (z-A_t)^{-1}A(z,t)^{N-1}\\ &= \sum_{j=0}^{N-1}(z-A_0)^{-1}A(z,t)^j + \left[(z-A_t)^{-1}-(z-A_0)^{-1}\right]A(z,t)^{N-1}.\end{aligned}$$

For each j, we then need to expand $A(z,t)^j$ to isolate the good Taylor expansion, and negligible terms. The computation is quite straightforward, but the notations are awful. To simplify them, let us denote by $\underline{\ell}$ a tuple $(\ell_1,\dots,\ell_k)$ of positive integers. Write also $l(\underline{\ell}) = k$ and $|\underline{\ell}| = \ell_1+\cdots+\ell_k$ and $\tilde{Q}_{\underline{\ell}} = Q_{\ell_1}(z-A_0)^{-1}\cdots Q_{\ell_k}(z-A_0)^{-1}$, and $\tilde{\Delta}_i(t) = \Delta_i(t)(z-A_0)^{-1}$.

Let us prove that, for any $j<N$,

$$\begin{aligned}A(z,t)^j = &\sum_{l(\underline{\ell})<j,\ j-l(\underline{\ell})<N-|\underline{\ell}|} t^{|\underline{\ell}|}A(z,t)^{j-l(\underline{\ell})-1}\tilde{\Delta}_{N-|\underline{\ell}|-(j-l(\underline{\ell})-1)}(t)\tilde{Q}_{\underline{\ell}}\\ &+\sum_{l(\underline{\ell})=j,\ 0<N-|\underline{\ell}|} t^{|\underline{\ell}|}\tilde{Q}_{\underline{\ell}}.\end{aligned} \tag{3.24}$$

We start from the following equality, valid for each $j\in\mathbb{N}$ and $a\le N$, which is a direct consequence of the definition of $\Delta_a(t)$:

$$A(z,t)^j = A(z,t)^{j-1}\tilde{\Delta}_a(t) + \sum_{\ell=1}^{a-1} t^\ell A(z,t)^{j-1}\tilde{Q}_\ell. \tag{3.25}$$

We can again iterate this equation. We will adjust the parameter a used during this iteration, as follows: we claim that, for all $1 \leq m \leq j$,

$$(3.26)\quad A(z,t)^j = \sum_{l(\underline{\ell})<m,\ j-l(\underline{\ell})<N-|\underline{\ell}|} t^{|\underline{\ell}|} A(z,t)^{j-l(\underline{\ell})-1} \tilde{\Delta}_{N-|\underline{\ell}|-(j-l(\underline{\ell})-1)}(t) \tilde{Q}_{\underline{\ell}} + \sum_{l(\underline{\ell})=m,\ j-l(\underline{\ell})<N-|\underline{\ell}|} t^{|\underline{\ell}|} A(z,t)^{j-m} \tilde{Q}_{\underline{\ell}}\,.$$

In fact, for $m = 1$ the above formula is just (3.25) for $a = N - j + 1$. Next, suppose (3.26) is true for some $m < j$, then the formula for $m + 1$ follows by substituting the last terms $A(z,t)^{j-m}$ using (3.25) for $a = N-|\underline{\ell}|-(j-l(\underline{\ell})-1)$. This proves (3.26) for any $m \leq j$. In particular, for $m = j$, we obtain (3.24).

The equations (3.23) and (3.24) sum up to

$$(3.27)\quad (z-A_t)^{-1} = R_N(t) + \left[(z-A_t)^{-1} - (z-A_0)^{-1}\right] A(z,t)^{N-1} + \sum_{j=0}^{N-1} \sum_{l(\underline{\ell})<j,\ j-l(\underline{\ell})<N-|\underline{\ell}|} t^{|\underline{\ell}|} (z-A_0)^{-1} A(z,t)^{j-l(\underline{\ell})-1} \tilde{\Delta}_{N-|\underline{\ell}|-(j-l(\underline{\ell})-1)}(t) \tilde{Q}_{\underline{\ell}}.$$

We will show that all the error terms are $O_{\mathcal{B}_N\to\mathcal{B}_0}(|t|^{\kappa b_0 + b(1,N)})$.

Fix j and $\underline{\ell}$ with $l(\underline{\ell}) < j,\ j - l(\underline{\ell}) < N - |\underline{\ell}|$. Let

$$(3.28)\quad F(t) = t^{|\underline{\ell}|} A(z,t)^{j-l(\underline{\ell})-1} \tilde{\Delta}_{N-|\underline{\ell}|-(j-l(\underline{\ell})-1)}(t) \tilde{Q}_{\underline{\ell}}.$$

We wish to show that

$$(3.29)\quad \left\|(z-A_0)^{-1} F(t)\right\|_{\mathcal{B}_N\to\mathcal{B}_0} \leq C|t|^{\kappa b_0 + b(1,N)}.$$

We have $\left\| |t|^{|\underline{\ell}|} \tilde{Q}_{\underline{\ell}} \right\|_{\mathcal{B}_N \to \mathcal{B}_{N-|\underline{\ell}|}} \leq C|t|^{|\underline{\ell}|} \leq C|t|^{b(N-|\underline{\ell}|,N)}$ by (3.16), while

$$(3.30)\quad \left\| \tilde{\Delta}_{N-|\underline{\ell}|-(j-l(\underline{\ell})-1)}(t) \right\|_{\mathcal{B}_{N-|\underline{\ell}|}\to\mathcal{B}_{j-l(\underline{\ell})-1}} \leq C|t|^{b(j-l(\underline{\ell})-1, N-|\underline{\ell}|)}$$

by (3.17), and $\left\|A(z,t)^{j-l(\underline{\ell})-1}\right\|_{\mathcal{B}_{j-l(\underline{\ell})-1}\to\mathcal{B}_0} \leq C|t|^{b(0,j-l(\underline{\ell})-1)}$ again by (3.17) applied $j - l(\underline{\ell}) - 1$ times, since $A(z,t) = \tilde{\Delta}_1(t)$. Multiplying these estimates gives

$$(3.31)\quad \|F(t)\|_{\mathcal{B}_N\to\mathcal{B}_0} \leq C|t|^{b(0,N)}.$$

Moreover, since $\tilde{\Delta}_k = \tilde{\Delta}_{k-1} - t^k \tilde{Q}_k$, the norm of $\tilde{\Delta}_k$ from $\mathcal{B}_j$ to $\mathcal{B}_{j-k+1}$ is bounded by $C|t|^{b(j-k+1,j)}$. In particular, the norm of $\tilde{\Delta}_{N-|\underline{\ell}|-(j-l(\underline{\ell})-1)}(t)$

from $\mathcal{B}_{N-|\underline{\ell}|}$ to $\mathcal{B}_{j-l(\underline{\ell})}$ is bounded by $C|t|^{b(j-l(\underline{\ell}),N-|\underline{\ell}|)}$. Together with the same arguments as above, we obtain

$$\|F(t)\|_{\mathcal{B}_N\to\mathcal{B}_1} \leq C|t|^{b(1,N)}. \tag{3.32}$$

The estimate (3.29) now follows from (3.31) and (3.32), as well as Lemma 3.4 for $\tau = |t|^{b_0}$.

We now turn to the term

$$\left[(z-A_t)^{-1} - (z-A_0)^{-1}\right] A(z,t)^{N-1} \tag{3.33}$$

of (3.27). As $\|A(z,t)\|_{\mathcal{B}_i\to\mathcal{B}_{i-1}} = O(|t|^{b_{i-1}})$, we have $\left\|A(z,t)^{N-1}\right\|_{\mathcal{B}_N\to\mathcal{B}_1} = O(|t|^{b(1,N)})$. With (3.18), this shows that (3.33) is $O_{\mathcal{B}_N\to\mathcal{B}_0}(|t|^{\kappa b_0+b(1,N)})$, concluding the proof. ∎

We will use the previous theorem in the following form:

Corollary 3.5: *Under the assumptions of the previous theorem, assume also that $M=1$ and that A_0 acting on each space $\mathcal{B}_j$ has a simple isolated eigenvalue at 1, with corresponding eigenfunction ξ_0. Then, for small enough t, A_t has a unique simple isolated eigenvalue $\lambda(t)$ close to 1.*

Let ν be a continuous linear form on $\mathcal{B}_0$ with $\nu(\xi_0)=1$. For small enough t, ν does not vanish on the eigenfunction of A_t for the eigenvalue $\lambda(t)$. It is therefore possible to define a normalized eigenfunction ξ_t satisfying $\nu(\xi_t)=1$.

Finally, there exist $u_1 \in \mathcal{B}_{N-1},\dots,u_{N-1}\in\mathcal{B}_1$ such that, for any $\epsilon>0$,

$$\left\|\xi_t - \xi_0 - \sum_{k=1}^{N-1} t^k u_k\right\|_{\mathcal{B}_0} = O(|t|^{b(0,N)-\epsilon}). \tag{3.34}$$

Proof. Let $c>0$ be small, and define an operator $P_t = \frac{1}{2\mathbf{i}\pi}\int_{|z-1|=c}(z-A_t)^{-1}\,\mathrm{d}z$. The operator P_0 is the spectral projection corresponding to the eigenvalue 1 of P_0. By Theorem 3.3, $\|P_t-P_0\|_{\mathcal{B}_N\to\mathcal{B}_0}$ converges to 0 when $t\to 0$. Therefore, the operator P_t is also a rank one projection for small enough t, corresponding to an eigenvalue $\lambda(t)$ of A_t. Let $\tilde{\xi}_t = P_t(u)$ for some fixed $u\in\mathcal{B}_N$ with $P_0(u)\neq 0$; $\tilde{\xi}_t$ is an eigenfunction of A_t for the eigenvalue $\lambda(t)$. Since $\|\tilde{\xi}_t-\tilde{\xi}_0\|_{\mathcal{B}_0}\to 0$, this eigenfunction satisfies $\nu(\tilde{\xi}_t)\neq 0$ for small enough t, and we can define a normalized eigenfunction $\xi_t = \tilde{\xi}_t/\nu(\tilde{\xi}_t)$.

For $1\leq k\leq N-1$, let

$$\tilde{u}_k = \sum_{\ell_1+\dots+\ell_j=k} \frac{1}{2\mathbf{i}\pi}\int_{|z-1|=c}(z-A_0)^{-1}Q_{\ell_1}\cdots(z-A_0)^{-1}Q_{\ell_j}(z-A_0)^{-1}u\,\mathrm{d}z.$$

It belongs to $\mathcal{B}_{N-k}$ by (3.16). Moreover, Theorem 3.3 yields

$$\left\| \tilde{\xi}_t - \tilde{\xi}_0 - \sum_{k=1}^{N-1} t^k \tilde{u}_k \right\|_{\mathcal{B}_0} \leq C|t|^{\kappa b_0 + b(1,N)}, \tag{3.35}$$

for $\kappa = \log((1-c)/\gamma)/\log(1/\gamma)$. Applying ν to this equation, we obtain that $\nu(\tilde{\xi}_t)$ admits an expansion $\nu(\tilde{\xi}_t) = \sum_{k=0}^{N-1} t^k \nu_k + O(|t|^{\kappa b_0 + b(1,N)})$. Hence, $\xi_t = \tilde{\xi}_t / \nu(\tilde{\xi}_t)$ also admits an expansion similar to (3.35).

This is almost the conclusion of the proof; we only have to see that the error term $O(|t|^{\kappa b_0 + b(1,N)})$ can be modified to be of the form $O(|t|^{b(0,N)-\epsilon})$ for any $\epsilon > 0$. This follows from the fact that c can be chosen arbitrarily small (by holomorphy, this does not change the projection P_t for small enough t, hence $\tilde{u}_k$ and u_k are also not modified). ∎

Remark 3.6: Corollary 3.5 states that the normalized eigenfunction ξ_t has a Taylor expansion of order $b(0,N) - \epsilon$ at 0. Under similar assumptions at every point of a neighborhood I of 0, we obtain that ξ_t has a Taylor expansion at every point of I. By a lemma of Campanato [Cam64], this implies that ξ_t is $C^{b(0,N)-\epsilon}$ on I, a result analogous to [HP08].

3.3. Definition of good Banach spaces. We now turn back to the dynamical setting: $T : X \to X$ is a mixing Gibbs–Markov map, and $f : X \to \mathbb{R}$ is a function satisfying $\sum m(a) Df(a)^\eta < \infty$, for which we want to prove an accurate characteristic expansion. To do that, we wish to apply Corollary 3.5 to a carefully chosen sequence of Banach spaces. We have currently at our disposal the spaces L^p (but the spectral properties of the transfer operator on these spaces are not good), and the space $\mathcal{L}$ (which is only a space, not a sequence of spaces). Our goal in this paragraph is to define a family of intermediate spaces between L^p and $\mathcal{L}$, which will be more suitable to apply Corollary 3.5.

For $1 \leq p \leq \infty$ and $s > 0$, let us define a Banach space $\mathcal{L}^{p,s}$ as follows: it is the space of measurable functions u such that, for any $k \in \mathbb{N}$, there exists a decomposition $u = v + w$ with $\|v\|_{\mathcal{L}} \leq Ce^k$ and $\|w\|_{L^p} \leq Ce^{-sk}$. The best such C is by definition the norm of u in $\mathcal{L}^{p,s}$. This Banach space is an *interpolation space* between $\mathcal{L}$ and L^p (see [BL76]).

Of course, $\mathcal{L}^{p,s}$ is included in L^p (simply use the decomposition for $k = 0$), and $\mathcal{L}^{p,s}$ is contained in $\mathcal{L}^{p',s'}$ when $p' \leq p$ and $s' \leq s$.

Let us check that the operators $\hat{T}$ and $\hat{T}_t$ enjoy good spectral properties when acting on $\mathcal{L}^{p,s}$. This will be a consequence of the fact that they have good properties when acting on $\mathcal{L}$, and are contractions when acting on L^p.

Lemma 3.7: *Let $1 \leq p \leq \infty$ and let $s > 0$. The operator $\hat{T}$ acts continuously on the space $\mathcal{L}^{p,s}$. Moreover, there exist $\gamma < 1$ and $C > 0$ such that*

$$\left\|\hat{T}^n u\right\|_{\mathcal{L}^{p,s}} \leq C\gamma^n \|u\|_{\mathcal{L}^{p,s}} + C\|u\|_{L^1}. \tag{3.36}$$

Proof. Let $\gamma_0 < 1$ be such that $\left\|\hat{T}^n u\right\|_{\mathcal{L}} \leq C\gamma_0^n \|u\|_{\mathcal{L}} + C\|u\|_{L^1}$.

For $n \in \mathbb{N}$, let A be the integer part of ϵn, for some $\epsilon > 0$ with $\gamma_0 e^{-\epsilon} < 1$. Let $u \in \mathcal{L}^{p,s}$; there exists a decomposition $u = v + w$ with $\|v\|_{\mathcal{L}} \leq e^{k+A}\|u\|_{\mathcal{L}^{p,s}}$ and $\|w\|_{L^p} \leq e^{-s(k+A)}\|u\|_{\mathcal{L}^{p,s}}$. Then

$$\begin{aligned}
\left\|\hat{T}^n v\right\|_{\mathcal{L}} &\leq C\gamma_0^n e^{k+A}\|u\|_{\mathcal{L}^{p,s}} + C\|v\|_{L^1} \\
&\leq C(\gamma_0^n e^A)e^k\|u\|_{\mathcal{L}^{p,s}} + C\|w\|_{L^1} + C\|u\|_{L^1} \\
&\leq C(\gamma_0^n e^A + e^{-s(k+A)})e^k\|u\|_{\mathcal{L}^{p,s}} + C\|u\|_{L^1} \\
&\leq e^k(C\gamma^n\|u\|_{\mathcal{L}^{p,s}} + C\|u\|_{L^1}),
\end{aligned}$$

for some $\gamma < 1$. Moreover,

$$\begin{aligned}
\left\|\hat{T}^n w\right\|_{L^p} &\leq \|w\|_{L^p} \leq Ce^{-sA}e^{-sk}\|u\|_{\mathcal{L}^{p,s}} \leq e^{-sk}(C\gamma^n\|u\|_{\mathcal{L}^{p,s}}) \\
&\leq e^{-sk}(C\gamma^n\|u\|_{\mathcal{L}^{p,s}} + C\|u\|_{L^1})
\end{aligned}$$

for some $\gamma < 1$.

Therefore, the decomposition of $\hat{T}^n u$ as $\hat{T}^n v + \hat{T}^n w$ shows that $\hat{T}^n u$ belongs to $\mathcal{L}^{p,s}$, and has a norm bounded by $C\gamma^n\|u\|_{\mathcal{L}^{p,s}} + C\|u\|_{L^1}$. ∎

Lemma 3.8: *For any $p \geq 1$ and $s > 0$, the inclusion of $\mathcal{L}^{p,s}$ in L^1 is compact.*

Proof. Let u_n be a sequence bounded by 1 in $\mathcal{L}^{p,s}$. Fix $k \in \mathbb{N}$, and let us decompose u_n as $v_n + w_n$ with $\|v_n\|_{\mathcal{L}} \leq e^k$ and $\|w_n\|_{L^p} \leq e^{-sk}$. Since the inclusion of $\mathcal{L}$ in L^1 is compact, there exists a subsequence $j(n)$ such that $v_{j(n)}$ converges in L^1. Therefore, $\limsup_{n,m\to\infty}\left\|u_{j(n)} - u_{j(m)}\right\|_{L^1} \leq 2e^{-sk}$. With a diagonal argument over k, we finally obtain a convergent subsequence of u_n. ∎

Corollary 3.9: *The transfer operator $\hat{T}$ acting on $\mathcal{L}^{p,s}$ has a simple eigenvalue at 1, and the rest of its spectrum is contained in a disk of radius < 1.*

Proof. Together with Hennion's Theorem [Hen93], the two previous lemmas ensure that the essential spectral radius of $\hat{T}$ acting on $\mathcal{L}^{p,s}$ is $\leq \gamma < 1$, i.e., the elements of the spectrum of $\hat{T}$ with modulus $> \gamma$ are isolated eigenvalues of finite multiplicity.

If u is an eigenfunction of $\hat{T}$ for an eigenvalue of modulus 1, then u belongs to L^1. Since $\hat{T}$ satisfies a Lasota–Yorke inequality (3.4) on the space $\mathcal{L}$, the theorem of Ionescu Tulcea and Marinescu [ITM50] implies that u belongs to $\mathcal{L}$. However, we know that $\hat{T}$ acting on $\mathcal{L}$ has a simple eigenvalue at 1, and no other eigenvalue of modulus 1. ∎

Lemma 3.10: *For any $p \geq 1$ and $s > 0$, the operator $\hat{T}_t$ acts continuously on $\mathcal{L}^{p,s}$ for small enough t. Moreover, $\|\hat{T}_t - \hat{T}\|_{\mathcal{L}^{p,s} \to \mathcal{L}^{p,s}}$ converges to 0 when $t \to 0$. Finally, if t is small enough, $\hat{T}_t$ satisfies a Lasota–Yorke type inequality*

$$\left\|\hat{T}_t^n u\right\|_{\mathcal{L}^{p,s}} \leq C\gamma^n \|u\|_{\mathcal{L}^{p,s}} + C \|u\|_{L^1}, \tag{3.37}$$

where $C > 0$ and $\gamma < 1$ are independent of t.

Proof. For any operator M sending $\mathcal{L}$ to $\mathcal{L}$ and L^p to L^p, then M sends $\mathcal{L}^{p,s}$ to $\mathcal{L}^{p,s}$ and, for any integer $A \geq 0$,

$$\|M\|_{\mathcal{L}^{p,s} \to \mathcal{L}^{p,s}} \leq \max(e^A \|M\|_{\mathcal{L} \to \mathcal{L}}, e^{-sA} \|M\|_{L^p \to L^p}). \tag{3.38}$$

This follows using the decomposition of $u \in \mathcal{L}^{p,s}$ as $v + w$ with $\|v\|_{\mathcal{L}} \leq e^{k+A} \|u\|_{\mathcal{L}^{p,s}}$ and $\|w\|_{L^p} \leq e^{-sk-sA} \|u\|_{\mathcal{L}^{p,s}}$.

By (3.5), $\|\hat{T}_t - \hat{T}\|_{\mathcal{L} \to \mathcal{L}}$ tends to 0, while $\|\hat{T}_t - \hat{T}\|_{L^p \to L^p}$ is uniformly bounded. Applying (3.38) to $M = \hat{T}_t - \hat{T}$ and e^A close to $\|\hat{T}_t - \hat{T}\|_{\mathcal{L} \to \mathcal{L}}^{-1/2}$, we obtain that $\|\hat{T}_t - \hat{T}\|_{\mathcal{L}^{p,s} \to \mathcal{L}^{p,s}}$ tends to zero.

By (3.36), we can fix $N > 0$, $\sigma < 1$ and $C > 0$ such that $\|\hat{T}^N u\|_{\mathcal{L}^{p,s}} \leq \sigma \|u\|_{\mathcal{L}^{p,s}} + C \|u\|_{L^1}$. Let $\sigma_1 \in (\sigma, 1)$. Since $\|\hat{T}_t - \hat{T}\|_{\mathcal{L}^{p,s} \to \mathcal{L}^{p,s}}$ tends to 0 when $t \to 0$, the previous equation gives, for small enough t,

$$\left\|\hat{T}_t^N u\right\|_{\mathcal{L}^{p,s}} \leq \sigma_1 \|u\|_{\mathcal{L}^{p,s}} + C \|u\|_{L^1}. \tag{3.39}$$

Iterating this equation, we get by induction over k

$$\left\|\hat{T}_t^{kN} u\right\|_{\mathcal{L}^{p,s}} \leq \sigma_1^k \|u\|_{\mathcal{L}^{p,s}} + C \sum_{j=0}^{k-1} \sigma_1^{k-1-j} \left\|\hat{T}_t^{jN} u\right\|_{L^1}. \tag{3.40}$$

Since $\hat{T}_t$ is a contraction on L^1, we obtain $\|\hat{T}_t^{kN} u\|_{\mathcal{L}^{p,s}} \leq \sigma_1^k \|u\|_{\mathcal{L}^{p,s}} + C' \|u\|_{L^1}$, for $C' = C\sum_{j=0}^{\infty} \sigma^j$. This proves (3.37) for n of the form kN, and the general case follows. ∎

3.4. Gaining δ in the integrability exponent. We wish to apply Corollary 3.5 to obtain the accurate characteristic expansion. This theorem involves an (arbitrarily) small loss of ϵ, that we will have to compensate at some point. In this paragraph, we show how a regularity assumption of the form $\sum m(a) Df(a)^\eta < \infty$ makes it possible to obtain a definite gain in the integrability exponent of some functions, which ultimately will compensate the aforementioned loss.

Lemma 3.11: *For any $\beta \in (0,1]$, there exists $\delta > 0$ with the following property. Let $f \in L^p$ (for some $p \in [1, 1/\beta]$) satisfy $\sum m(a) Df(a)^\beta < \infty$. Let $c \in [\beta, p]$, and consider a function u such that $|u| \leq |f|^c$, and, for all $a \in \alpha$,*

$$Du(a) \leq \begin{cases} Df(a) & \text{if } c \leq 1, \\ Df(a) \|1_a f\|_{L^\infty}^{c-1} & \text{if } c > 1. \end{cases} \tag{3.41}$$

Let q, r be positive numbers (possibly $q = \infty$) such that $1/r = 1/(p/c) + 1/q$, and $r \geq 1 + \beta$. Then the operator $v \mapsto \hat{T}(uv)$ maps L^q to $L^{r+\delta}$ (and its norm is bounded only in terms of f and β).

Since $u \in L^{p/c}$, the Hölder inequality shows that the operator $v \mapsto \hat{T}(uv)$ maps L^q to L^r. The lemma asserts that there is in fact a small gain of δ in the integrability exponent, due to the regularity property $\sum m(a) Df(a)^\beta < \infty$. Moreover, the gain is uniform over the parameters if r stays away from 1.

Proof. We will show that, under the assumptions of the lemma, the operator $v \mapsto \hat{T}(uv)$ maps L^q to $L^{\tilde{r}}$, for

$$\tilde{r} = \frac{pq/c - \beta^2 q}{p/c + q - \beta^2 q}.$$

Since $\tilde{r} - r$ is uniformly bounded from below when the parameters vary according to the conditions of the lemma, this will conclude the proof.[2]

[2] Indeed,

$$\tilde{r} - r = \frac{\beta^2 q\,(pq/c - p/c - q)}{(p/c+q)(p/c + q - \beta^2 q)}. \tag{3.42}$$

Let us first show that

$$\sum_{a\in\alpha} m(a)\,\|f1_a\|_{L^\infty}^{\beta} < \infty. \tag{3.44}$$

For $x, y \in a$, we have $|f(x)| \leq |f(y)| + Df(a)$. Integrating over y, we get $|f(x)| \leq \frac{1}{m(a)}\int_a |f| + Df(a)$. Together with the inequality $(t'+t)^\beta \leq 1+t'+t^\beta$, valid for any $t', t \geq 0$, we obtain

$$\sum_{a\in\alpha} m(a)\,\|f1_a\|_{L^\infty}^{\beta} \leq \sum_{a\in\alpha} m(a)\left(1+\frac{1}{m(a)}\int_a |f|\right) + \sum_{a\in\alpha} m(a) Df(a)^\beta.$$

These sums are finite, concluding the proof of (3.44).

Let $\tilde\beta = \beta^2$; we will now show that

$$\sum_{a\in\alpha} m(a)\,\|u1_a\|_{L^\infty}^{\tilde\beta} < \infty. \tag{3.45}$$

By the previous argument, it is sufficient to show that $\sum_{a\in\alpha} m(a)Du(a)^{\tilde\beta}$ is finite. If $c \leq 1$, $Du(a)^{\tilde\beta} \leq Df(a)^{\tilde\beta} \leq \max(1, Df(a)^\beta)$, and the result follows. If $c > 1$,

$$\begin{aligned} Du(a)^{\tilde\beta} &\leq Df(a)^{\tilde\beta}\,\|1_a f\|_{L^\infty}^{(c-1)\tilde\beta} \leq \max(Df(a), \|1_a f\|_{L^\infty})^{c\tilde\beta} \\ &\leq Df(a)^{c\tilde\beta} + \|1_a f\|_{L^\infty}^{c\tilde\beta}. \end{aligned}$$

Since $c\tilde\beta \leq \beta$, (3.44) shows that $\sum m(a)Du(a)^{\tilde\beta} < \infty$, concluding the proof of (3.45).

By (3.1) and (3.2), $\hat T(|u|^{\tilde\beta})$ is bounded by $\sum_{a\in\alpha} m(a)\,\|u1_a\|_{L^\infty}^{\tilde\beta}$, which is finite. Hence, $\hat T(|u|^{\tilde\beta})$ is a bounded function.

Let us now estimate $\hat T(uv)$ for $v \in L^q$. Let $\rho = \tilde r/(\tilde r-1)$, so that $1/\rho+1/\tilde r = 1$. We have

$$\hat T(|uv|) = \hat T(|u|^{\tilde\beta/\rho}|u|^{1-\tilde\beta/\rho}|v|) \leq \hat T(|u|^{\tilde\beta})^{1/\rho}\hat T(|u|^{\tilde r(1-\tilde\beta/\rho)}|v|^{\tilde r})^{1/\tilde r}. \tag{3.46}$$

Since $\hat T(|u|^{\tilde\beta})$ is bounded, we obtain

$$\left\|\hat T(uv)\right\|_{L^{\tilde r}} \leq C\left(\int \hat T(|u|^{\tilde r(1-\tilde\beta/\rho)}|v|^{\tilde r})\right)^{1/\tilde r} = C\left(\int |u|^{\tilde r(1-\tilde\beta/\rho)}|v|^{\tilde r}\right)^{1/\tilde r}.$$

Since $r \geq 1+\beta$, we have $pq/c \geq (1+\beta)(p/c+q)$. Therefore, the second term of the numerator of (3.42) is at least $\beta(p/c+q)$. Simplifying with the denominator, we get

$$\tilde r - r \geq \frac{\beta^3}{p/qc+1-\beta^2} \geq \frac{\beta^3}{\beta^{-2}+1-\beta^2}. \tag{3.43}$$

Let s and t be such that $1/s+1/t=1$ and $t\tilde{r}=q$, i.e., $t=q/\tilde{r}$ and $s=q/(q-\tilde{r})$. The Hölder inequality gives

$$\int |u|^{\tilde{r}(1-\tilde{\beta}/\rho)}|v|^{\tilde{r}} \leq \left(\int |u|^{\tilde{r}(1-\tilde{\beta}/\rho)s}\right)^{1/s}\left(\int |v|^{\tilde{r}t}\right)^{1/t}. \tag{3.47}$$

The choice of $\tilde{r}$ above ensures that $\tilde{r}(1-\tilde{\beta}/\rho)s = p/c$. Hence, the integral involving u is finite, since $u \in L^{p/c}$. We obtain $\|\hat{T}(uv)\|_{L^{\tilde{r}}} \leq C\|v\|_{L^q}$, as required. ∎

Lemma 3.12: *For any $\beta \in (0,1]$, there exists $\delta>0$ with the following property. Let $f \in L^p$ (for some $p\in[1,1/\beta]$) satisfy $\sum m(a)Df(a)^\beta < \infty$. Let $c\in[\beta,p]$, and consider a function u such that $|u| \leq |f|^c$, and, for all $a\in\alpha$,*

$$Du(a) \leq \begin{cases} Df(a) & \text{if } c\leq 1,\\ Df(a)\,\|1_a f\|_{L^\infty}^{c-1} & \text{if } c>1. \end{cases} \tag{3.48}$$

Let q, r be positive numbers (possibly $q=\infty$) such that $1/r = 1/(p/c)+1/q$, and $r\geq 1+\beta$. Then, for any $s>0$, there exists $s'=s'(f,\beta,s)$ such that the operator $v\mapsto \hat{T}(uv)$ maps $\mathcal{L}^{q,s}$ to $\mathcal{L}^{r+\delta,s'}$ (and its norm is bounded only in terms of f,β,s).

Proof. Let δ_0 be the value of δ given by Lemma 3.11 for $\beta/2$ instead of β. We will prove that the lemma holds for $\delta=\delta_0/2$.

For $K\geq 1$, denote by $A(K)$ the union of the elements $a\in\alpha$ such that $Df(a)+\|1_af\|_{L^\infty}\leq K$, and let $B(K)$ be its complement. The finiteness of the sum $\sum_{a\in\alpha} m(a)(Df(a)^\beta + \|1_a f\|_{L^\infty}^\beta)$ (which has been proved in (3.44)) implies that there exists C such that

$$m(B(K)) \leq CK^{-\beta}. \tag{3.49}$$

Moreover, let $d=\max(c,1)$; then u is bounded by K^d on $A(K)$, and its Lipschitz constant is also bounded by K^d. Therefore,

$$\left\|1_{A(K)}u\right\|_{\mathcal{L}} \leq CK^d. \tag{3.50}$$

Take $v\in\mathcal{L}^{q,s}$ bounded by 1, and $k\in\mathbb{N}$. By definition of $\mathcal{L}^{q,s}$, we can write $v=w+w'$ with $\|w\|_{\mathcal{L}}\leq e^k$ and $\|w'\|_{L^q}\leq e^{-sk}$. For any $K\geq 1$, we obtain a decomposition of $\hat{T}(uv)$ as the sum of $\hat{T}(1_{A(K)}uw)$ and $\hat{T}(1_{B(K)}uw+uw')$. We claim that

$$\left\|\hat{T}(1_{A(K)}uw)\right\|_{\mathcal{L}} \leq CK^d e^k \tag{3.51}$$

and, for some $\epsilon > 0$,

$$\left\| \hat{T}(1_{B(K)}uw + uw') \right\|_{L^{r+\delta_0/2}} \leq Ce^{-sk} + Ce^k K^{-\epsilon}. \tag{3.52}$$

This concludes the proof of the lemma, for $\delta = \delta_0/2$ and $s' = s/(1+d(1+s)/\epsilon)$. Indeed, for $K = \exp((1+s)k/\epsilon)$, the bound in (3.51) becomes $Ce^{sk/s'}$, and the bound in (3.52) becomes Ce^{-sk}. Taking k' close to sk/s', we have obtained a decomposition of $\hat{T}(uv)$ as a sum $\tilde{w} + \tilde{w}'$ with $\|\tilde{w}\|_{\mathcal{L}} \leq Ce^{k'}$ and $\|\tilde{w}'\|_{L^{r+\delta_0/2}} \leq Ce^{-s'k'}$, as desired.

It remains to prove (3.51) and (3.52). The former follows from the inequality $\|zz'\|_{\mathcal{L}} \leq C\|z\|_{\mathcal{L}} \|z'\|_{\mathcal{L}}$, applied to the functions $z = 1_{A(K)}u$ (whose norm is bounded by (3.50)), and $z' = w$ (whose norm is at most e^k).

We turn to (3.52). First, by Lemma 3.11,

$$\left\| \hat{T}(uw') \right\|_{L^{r+\delta_0/2}} \leq \left\| \hat{T}(uw') \right\|_{L^{r+\delta_0}} \leq C \|w'\|_{L^q}, \tag{3.53}$$

which is at most e^{-sk}. Let then Q be large enough, and let r' be such that $1/r' = 1/(p/c) + 1/Q$. Then

$$r - r' = rr' \left(\frac{1}{r'} - \frac{1}{r} \right) = rr' \left(\frac{1}{Q} - \frac{1}{q} \right) \leq \frac{rr'}{Q}. \tag{3.54}$$

Moreover, $1/r \geq 1/(p/c) \geq \beta^2$, and $1/r' \geq \beta^2$ as well. Hence, $r - r' \leq \beta^{-4}/Q$. Choosing Q large enough, we can ensure $r - r' \leq \delta_0/2$. Therefore,

$$\begin{aligned} \left\| \hat{T}(1_{B(K)}uw) \right\|_{L^{r+\delta_0/2}} &\leq \|w\|_{L^\infty} \left\| \hat{T}(1_{B(K)}|u|) \right\|_{L^{r+\delta_0/2}} \\ &\leq \|w\|_{L^\infty} \left\| \hat{T}(1_{B(K)}|u|) \right\|_{L^{r'+\delta_0}}. \end{aligned}$$

Since $1/(p/c) \leq 1/r \leq 1/(1+\beta)$, we have $1/r' \leq 1/(1+\beta) + 1/Q$, which is at most $1/(1+\beta/2)$ if Q is large enough. Thanks to the definition of δ_0 above, we can therefore apply Lemma 3.11 to the function $v = 1_{B(K)}$ and the parameters $r', Q, \beta/2$, to obtain

$$\left\| \hat{T}(1_{B(K)}|u|) \right\|_{L^{r'+\delta_0}} \leq C \left\| 1_{B(K)} \right\|_{L^Q}. \tag{3.55}$$

Since $\|w\|_{L^\infty} \leq e^k$ and $\left\| 1_{B(K)} \right\|_{L^Q} \leq CK^{-\beta/Q}$ by (3.49), this proves (3.52) for $\epsilon = \beta/Q$. ∎

3.5. Accurate characteristic expansions for integrable functions. We will now prove that a function f satisfying $\sum m(a)Df(a)^\eta < \infty$ admits an admissible characteristic expansion. By Lemma 3.2, it is sufficient to treat the case $f \in L^{1+\eta/2}$. We will give very precise asymptotics of the eigenvalue $\lambda(t)$ of the transfer operator, yielding also other limit theorems in the L^2 case.

Theorem 3.13: *Let $\eta \in (0,1]$. There exists a function $\epsilon : (1,\infty) \to \mathbb{R}_+^*$, bounded away from zero on compact subsets of $(1,\infty)$, with the following property.*

Let f satisfy $\sum m(a)Df(a)^\eta < \infty$, and $f \in L^p$ for some $p > 1$. Then there exist complex numbers c_i (for $1 \le i < p + \epsilon(p)$) such that

$$\lambda(t) = E(e^{\mathbf{i}tf}) + \sum_{2 \le i < p+\epsilon(p)} c_i t^i + O(|t|^{p+\epsilon(p)}). \tag{3.56}$$

This theorem contains the characteristic expansion of $f \in L^p$ for $p > 1$:

Corollary 3.14: *Let $f \in L^{1+\eta/2}$; then f admits an accurate characteristic expansion.*

Proof. If $f \in L^2$, then (3.56) for $p = 2$ becomes $\lambda(t) = 1 + \mathbf{i}tE(f) - ct^2/2 + o(t^2)$, for some $c \in \mathbb{C}$. This is the desired characteristic expansion.

Assume now $f \notin L^2$. Let $\epsilon > 0$ be the infimum of $\epsilon(p)$ for $p \in [1+\eta/2, 2]$. Let $p \ge 1 + \eta/2$ be such that $f \in L^p$ and $f \notin L^{p+\epsilon/2}$. Then (3.56) gives $\lambda(t) = E(e^{\mathbf{i}tf}) + ct^2 + O(|t|^{p+\epsilon})$, which is accurate. ∎

Together with Lemma 3.2, this concludes the proof of Proposition 1.10.

Theorem 3.13 also contains much more information, in particular in the L^2 case. We will describe in Appendix A another consequence of this very precise expansion of the eigenvalue $\lambda(t)$, on the speed of convergence in the central limit theorem. What is remarkable in that theorem is that the regularity assumption on the function need not be increased to get finer results, $\sum m(a)Df(a)^\eta < \infty$ is always sufficient: the only additional conditions are moment conditions.

Remark 3.15: For $p < 2$, Theorem 3.13 can be proved using only the theorem of Keller and Liverani in [KL99], instead of its extension to several derivatives given in Paragraph 3.2 (but the resulting bound ϵ tends to 0 when p tends to 2): in the forthcoming proof, there is no derivative involved for $p < 2$. This gives a more elementary proof of the accurate characteristic expansion for functions

f not belonging to L^q for some $q < 2$, but the general case (functions in L^p for every $p < 2$) requires the full power of Theorem 3.3.

We will need the following elementary lemma.

LEMMA 3.16: *For $j \geq 1$, define a function $F_j : \mathbb{R} \to \mathbb{C}$ by*

$$F_j(x) = e^{\mathbf{i}x} - \sum_{k=0}^{j-1} (\mathbf{i}x)^k / k!. \tag{3.57}$$

Let also $b \in (0, 1]$. For $j \geq 1$ and $x \in \mathbb{R}$, $|F_j(x)| \leq 2|x|^{j-1+b}$. Moreover, for $j \geq 2$ and $x, y \in \mathbb{R}$, $|F_j(x) - F_j(y)| \leq 2|x - y| \cdot \max(|x|, |y|)^{j-2+b}$.

Proof. Let (A_j) denote the property "for all $x \in \mathbb{R}$, $|F_j(x)| \leq 2|x|^{j-1+b}$" and (B_j) the property "for all x, y, $|F_j(x) - F_j(y)| \leq 2|x - y| \cdot \max(|x|, |y|)^{j-2+b}$". We claim that (A_j) holds for $j \geq 1$, and (B_j) holds for $j \geq 2$.

First, (A_1) holds trivially. Moreover, if (B_j) holds, then (A_j) holds by taking $y = 0$. Hence, it is sufficient to prove that (A_j) implies (B_{j+1}) to conclude by induction. Assume (A_j). Since $F'_{j+1} = \mathbf{i}F_j$, we have

$$\begin{aligned} |F_{j+1}(x) - F_{j+1}(y)| &\leq |x - y| \sup_{z \in [x,y]} |F'_{j+1}(z)| \leq |x - y| \sup_{z \in [x,y]} 2|z|^{j-1+b} \\ &\leq 2|x - y| \max(|x|, |y|)^{j-1+b}. \end{aligned}$$

This proves (B_{j+1}), as desired. ∎

Proof of Theorem 3.13. Fix once and for all $\eta \in (0, 1]$. Let us fix $A > 1$; we will prove the theorem for $p \in [1 + 1/A, A]$.

The quantity $\frac{p}{1+1/5A} - \frac{p}{1+1/2A}$ is bounded from below, uniformly for $p \in [1 + 1/A, A]$. Therefore, there exists $\delta_p \in [1/5A, 1/2A]$ such that the distance from $p/(1 + \delta_p)$ to the integers is $\geq \overline{\delta}$, for some $\overline{\delta} > 0$. Let us fix such a δ_p. Let $N \geq 2$ be the integer such that $N > p/(1 + \delta_p) > N - 1$; write $p/(1 + \delta_p) = N - 1 + b_0$ for some $b_0 \in [\overline{\delta}, 1 - \overline{\delta}]$. Let $b_1 = \cdots = b_{N-1} = 1$. Define numbers $p_0, \ldots, p_N$ in $[1 + \delta_p, \infty]$ by $p_0 = 1 + \delta_p$ and, for $i \geq 1$, $p_i = p/(N - i)$. Define also operators Q_j by $Q_j(v) = \frac{\mathbf{i}^j}{j!} \hat{T}(f^j v)$, and let $\Delta_j(t) = \hat{T}_t - \hat{T} - \sum_{k=1}^{j-1} t^k Q_k$. Let $\tilde{\mathcal{B}}_j = L^{p_j}$.

We claim that the assumptions (3.16) and (3.17) are satisfied for the spaces $\tilde{\mathcal{B}}_j$. Indeed, the choices of b_0 and the p_is ensure that, for $0 \leq i < j \leq N$,

$$\frac{1}{p_i} = \frac{1}{p_j} + \frac{b(i, j)}{p}. \tag{3.58}$$

Therefore, if $u \in L^{p/b(i,j)}$ and $v \in L^{p_j}$, then $uv \in L^{p_i}$. Since $f^{j-i} \in L^{p/(j-i)}$, this shows that Q_{j-i} sends $\tilde{\mathcal{B}}^j$ to $\tilde{\mathcal{B}}^i$ if $i > 0$.

By Lemma 3.16, for any $n \geq 1$ and $b > 0$, $|e^{\mathbf{i}x} - \sum_{k=0}^{n-1} \frac{\mathbf{i}^k}{k!} x^k| \leq 2|x|^{n-1+b}$. Therefore, $|\Delta_{j-i}(t)v| \leq 2\hat{T}(|tf|^{j-i-1+b}|v|)$. Taking $b = b_i$, we obtain

$$|\Delta_{j-i}(t)v| \leq 2|t|^{b(i,j)}\hat{T}(|f|^{b(i,j)}|v|). \tag{3.59}$$

Since $|f|^{b(i,j)}$ belongs to $L^{p/b(i,j)}$, this shows, thanks to (3.58), that $\Delta_{j-i}(t)$ sends $\tilde{\mathcal{B}}_j$ to $\tilde{\mathcal{B}}_i$ with a norm at most $C|t|^{b(i,j)}$. This is (3.17).

Unfortunately, the spaces L^{p_j} do not satisfy a Lasota–Yorke type inequality (3.15). Moreover, we would like to gain a little bit on the integrability exponent. Therefore, we will rather use spaces $\mathcal{L}^{q,s}$ instead of spaces L^q. To check the assumptions (3.16) and (3.17), we will apply Lemma 3.12 for some small enough $\beta \in (0, \eta]$ depending only on A.

The assumptions of this lemma are satisfied for the operator Q_{j-i} $(1 \leq i < j \leq N)$, with $q = p_j$, $r = p_i$ and $c = b(i,j)$ (since f^{j-i} is indeed bounded by $|f|^{j-i}$, and $D(f^{j-i})(a) \leq CDf(a) \|1_a f\|_{L^\infty}^{j-i-1}$). We now turn, for $0 \leq i < j \leq N$, to the operators $\Delta_{j-i}(t)$. Once again, we take $q = p_j$, $r = p_i$ and $c = b(i,j)$. Let us show that the assumptions of Lemma 3.12 are satisfied. First, if β is small enough, then $r = p_i$ is larger than $1+\beta$ (since we have chosen $p_0 = 1+\delta_p$ with $\delta_p \geq 1/5A$), and $c = b(i,j)$ is larger than β (since $b_0 \geq \overline{\delta}$ by the good choice of δ_p). Let us define a function $f_{j-i}(t) = e^{\mathbf{i}tf} - \sum_{k=0}^{j-i-1} \frac{(\mathbf{i}tf)^k}{k!}$, so that $\Delta_{j-i}(t)v = \hat{T}(f_{j-i}(t)v)$. The following lemma shows that $f_{j-i}(t)$ is well behaved, which is the last assumption of Lemma 3.12 we have to check.

Lemma 3.17: *For any $0 < b \leq 1$ and $j \geq 1$, the function $u_j(t) = f_j(t)/(2|t|^{j-1+b})$ satisfies $|u_j| \leq |f|^{j-1+b}$ and, for all $a \in \alpha$,*

$$Du_j(t)(a) \leq \begin{cases} Df(a) & \text{if } j-1+b \leq 1, \\ Df(a)\,\|1_a f\|_{L^\infty}^{j-2+b} & \text{if } j-1+b > 1. \end{cases} \tag{3.60}$$

Proof. We have $f_j(t) = F_j(tf)$, where F_j is defined in Lemma 3.16. Therefore, this lemma yields $|f_j(t)| \leq 2|tf|^{j-1+b}$ as desired. If $j = 1$, $f_j(t) = e^{\mathbf{i}tf} - 1$, hence (3.60) follows easily. Assume now $j \geq 2$. For any points x, y in the same element a of the partition α,

$$\begin{aligned}|f_j(t)(x) - f_j(t)(y)| &= |F_j(tf(x)) - F_j(tf(y))| \\ &\leq 2|tf(x) - tf(y)| \max(|tf(x)|, |tf(y)|)^{j-2+b} \\ &\leq 2|t|^{j-1+b} Df(a) d(x,y) \left\| 1_a f \right\|_{L^\infty}^{j-2+b}.\end{aligned}$$

This proves (3.60) in this case. ∎

Let $\delta > 0$ be given by Lemma 3.12 for the value of β we constructed above. Decreasing δ if necessary, we can assume $\delta \leq 1/2A$. Let also $s_N = 1$. Lemma 3.12 (applied to the operators Q_1 and $\Delta_1(t)$, on the space $\mathcal{L}^{p_N, s_N}$) provides us with $s_{N-1} = s'$ such that $\|Q_1\|_{\mathcal{L}^{p_N,s_N} \to \mathcal{L}^{p_{N-1}+\delta, s_{N-1}}}$ is finite, and

$$\|\Delta_1(t)\|_{\mathcal{L}^{p_N,s_N} \to \mathcal{L}^{p_{N-1}+\delta, s_{N-1}}} = O(|t|^{b(N-1,N)}). \tag{3.61}$$

Continuing inductively this process, we obtain a sequence $s_N, s_{N-1}, \dots, s_0$ such that, for any $1 \leq i < j \leq N$, the operator Q_{j-i} maps continuously $\mathcal{L}^{p_j,s_j}$ to $\mathcal{L}^{p_i+\delta,s_i}$, and such that, for any $0 \leq i < j \leq N$, the operator $\Delta_{j-i}(t)$ maps continuously $\mathcal{L}^{p_j,s_j}$ to $\mathcal{L}^{p_i+\delta,s_i}$, with a norm at most $C|t|^{b(i,j)}$.

Define a space $\mathcal{B}_i = \mathcal{L}^{p_i+\delta,s_i}$. Since $\mathcal{B}_i$ is continuously contained in $\mathcal{L}^{p_i,s_i}$, we have just proved that the assumptions (3.16) and (3.17) of Theorem 3.3 are satisfied. Moreover, (3.14) and (3.15) for $M = 1$ follow from Lemmas 3.7 and 3.10. Therefore, Corollary 3.5 applies. Since $\mathcal{B}_i$ is included in $L^{p_i+\delta}$, we obtain in particular the following: there exist $u_1 \in L^{p_{N-1}+\delta}, \dots, u_{N-1} \in L^{p_1+\delta}$ such that the normalized eigenfunction ξ_t of $\hat{T}_t$ satisfies

$$\left\| \xi_t - 1 - \sum_{k=1}^{N-1} t^k u_k \right\|_{L^{p_0+\delta}} = O(|t|^{b(0,N)-\epsilon}), \tag{3.62}$$

for any $\epsilon > 0$.

Let us now estimate the eigenvalue $\lambda(t)$ of $\hat{T}_t$ using this estimate. Let us write $\xi_t - 1 = \sum_{k=1}^{N-1} t^k u_k + r_t$, where r_t is an error term controlled by (3.62). By (3.11),

$$\begin{aligned}\lambda(t) &= E(e^{\mathbf{i}tf}) + \int (e^{\mathbf{i}tf} - 1)(\xi_t - 1) \\ &= E(e^{\mathbf{i}tf}) + \sum_{k=1}^{N-1} t^k \int (e^{\mathbf{i}tf} - 1) u_k + \int (e^{\mathbf{i}tf} - 1) r_t.\end{aligned} \tag{3.63}$$

Let us first estimate $\int (e^{\mathbf{i}tf}-1)r_t$. We have $p_0 = 1+\delta_p$ and $b(0,N) = p/(1+\delta_p)$. Let q be such that $1/(p_0+\delta)+1/q=1$, i.e., $q=(1+\delta_p+\delta)/(\delta_p+\delta)$. Since $\delta \leq 1/2A$ and $\delta_p \leq 1/2A$, we obtain $q \geq A$. In particular, $q \geq p$. Therefore, $|e^{\mathbf{i}x}-1| \leq 2|x|^{p/q}$ for any real x. This yields

$$\left\| e^{\mathbf{i}tf}-1 \right\|_{L^q} \leq \left(\int |e^{\mathbf{i}tf}-1|^q \right)^{1/q} \leq C|t|^{p/q}. \tag{3.64}$$

Hence,

$$\left| \int (e^{\mathbf{i}tf}-1) r_t \right| \leq \left\| e^{\mathbf{i}tf}-1 \right\|_{L^q} \|r_t\|_{L^{p_0+\delta}} \leq C|t|^{p/q+p/(1+\delta_p)-\epsilon}. \tag{3.65}$$

Moreover,

$$\frac{p}{q}+\frac{p}{1+\delta_p}-\epsilon = p\left(1-\frac{1}{1+\delta_p+\delta}+\frac{1}{1+\delta_p}\right)-\epsilon. \tag{3.66}$$

Since δ is positive, this quantity is larger than p if ϵ is small enough. Hence, (3.65) is of the form $O(|t|^{p+\epsilon'})$ for some $\epsilon'>0$. This is compatible with (3.56).

We now turn to the terms $t^k \int (e^{\mathbf{i}tf}-1)u_k$ in (3.63), for $0 \leq k \leq N-1$. The function u_k belongs to $L^{p/k+\delta}$. Let q be such that $1/q+1/(p/k+\delta)=1$. Let also $c>0$ satisfy $qc=p$. Then $e^{\mathbf{i}tf} = \sum_{0\leq j<c} (\mathbf{i}tf)^j/j! + r_{c,t}$, where $|r_{c,t}| \leq 2|t|^c|f|^c$ by Lemma 3.16. To conclude the proof, it is sufficient to show that $t^k \int r_{c,t}u_k = O(|t|^{p+\epsilon'})$ for some $\epsilon'>0$, since the terms coming from the integrals $t^k \int (\mathbf{i}tf)^j/j! \cdot u_k$ will contribute to the polynomial in (3.56). We have

$$\begin{aligned}\left| t^k \int r_{c,t}u_k \right| &\leq |t|^k \left\|r_{c,t}\right\|_{L^q} \left\|u_k\right\|_{L^{p/k+\delta}} \leq C|t|^k \left(\int |r_{c,t}|^q \right)^{1/q} \\ &\leq C|t|^{k+c} \left(\int |f|^p \right)^{1/q}.\end{aligned}$$

Finally, $k+c = k+p-k/(1+k\delta/p)$ is strictly larger than p, since $\delta>0$. ∎

Remark 3.18: When $f \in L^p$, $p>1$, the function $\mu(t) = \int P_t 1$ appearing in the characteristic expansion (3.9) of f also satisfies an expansion

$$\mu(t) = 1 + \sum_{1\leq i<p} d_i t^i + O(|t|^{p-\epsilon}), \tag{3.67}$$

for any $\epsilon>0$. This follows from a similar (but easier) argument, where one does not need to use the gain in the exponent from Lemma 3.12. This expansion is not as strong as the expansion of $\lambda(t)$ (it does not reach the precision $O(|t|^p)$, while Theorem 3.13 gets beyond it). The reason for this difference is that $\mu(t)$

is only expressed in spectral terms (and Theorem 3.3 therefore gives a small loss in the exponent), while for $\lambda(t)$ one can take advantage of the formula (3.11).

3.6. Last details in the L^2 case. In this paragraph, we conclude the proof of Theorem 1.5. By Proposition 1.10 and Theorem 1.9, we only have to identify the variance σ^2 when $f \in L^2$, and to strengthen the conclusion of Theorem 1.9 in the $\sigma^2 = 0$ case.

Lemma 3.19: *Assume $f \in L^2$, and write $\tilde{f} = f - \int f$. Then the asymptotic expansion of $\lambda(t)$ given by Theorem 3.13 is*

$$\lambda(t) = 1 + \mathbf{i}tE(f) - (\sigma^2 + E(f)^2)t^2/2 + o(t^2), \tag{3.68}$$

where $\sigma^2 = \int \tilde{f}^2 + 2\sum_{k=1}^{\infty} \int \tilde{f} \cdot \tilde{f} \circ T^k$ (the series converges exponentially fast).

Proof. In the expansion (3.56) of $\lambda(t)$, the term for $i = 2$ comes only, in the proof, from the integral $\int \mathbf{i}tfu_1$, where

$$u_1 = \frac{1}{2\mathbf{i}\pi} \int_{|z-1|=c} (z - \hat{T})^{-1} Q_1 (z - \hat{T})^{-1} 1 \,\mathrm{d}z, \tag{3.69}$$

where $Q_1(v) = \hat{T}(\mathbf{i}fv)$ (and the integral is converging in a space $\mathcal{L}^{2+\delta,s}$ for some $\delta > 0$ and $s > 0$). Let us identify u_1. We have $(z - \hat{T})^{-1}1 = 1/(z-1)$. Moreover, if E is the space of constant functions and F the space of functions with vanishing integral, then $(z - \hat{T})^{-1}$ is the multiplication by $1/(z-1)$ on E, while $(z - \hat{T})^{-1}v = \sum_{k=0}^{\infty} z^{-k-1}\hat{T}^k v$ for $v \in F$ (the series converging exponentially fast in $\mathcal{L}^{2+\delta,s}$, and in particular in L^2). Writing $\hat{T}f$ as $(\int f) + \hat{T}\tilde{f} \in E \oplus F$, we obtain

$$u_1 = \frac{1}{2\mathbf{i}\pi} \int_{|z-1|=c} \frac{\mathbf{i}\int f}{(z-1)^2} \,\mathrm{d}z + \sum_{k=0}^{\infty} \frac{1}{2\mathbf{i}\pi} \int_{|z-1|=c} \frac{z^{-k-1}}{z-1} \mathbf{i}\hat{T}^{k+1}\tilde{f} \,\mathrm{d}z. \tag{3.70}$$

Since $1/(z-1)^2$ has a vanishing residue at $z = 1$, while $z^{-k-1}/(z-1)$ has a residue equal to 1, this gives

$$u_1 = \mathbf{i} \sum_{k=0}^{\infty} \hat{T}^{k+1}\tilde{f}. \tag{3.71}$$

We obtain from (3.56)

$$\begin{aligned}\lambda(t) &= E(e^{\mathbf{i}tf}) - t^2 \sum_{k=1}^{\infty} \int f\hat{T}^k \tilde{f} + O(|t|^{2+\epsilon}) \\ &= 1 + \mathbf{i}tE(f) - t^2 \int f^2/2 - t^2 \sum_{k=1}^{\infty} \int \tilde{f} \cdot \tilde{f} \circ T^k + o(t^2) \\ &= 1 + \mathbf{i}tE(f) - (\sigma^2 + E(f)^2)t^2/2 + o(t^2). \qquad \blacksquare\end{aligned}$$

To conclude, it is sufficient to prove that, if σ^2 vanishes, then f is a bounded coboundary. A similar result is proved in [AD01b, Corollary 2.3], and we will essentially reproduce the same argument for completeness.

LEMMA 3.20: *Assume $f \in L^2$ is such that σ^2 (given by Lemma 3.19) vanishes. Then there exist a bounded function u and a real c such that $f = u - u \circ T + c$.*

Proof. Replacing f with $\tilde{f} = f - \int f$, we can assume without loss of generality that $\int f = 0$.

The exponential convergence of $\int f \cdot f \circ T^k$ to 0 ensures that $\int (S_n f)^2 = n\sigma^2 + O(1)$. Therefore, if $\sigma^2 = 0$, then $S_n f$ is bounded in L^2. By Leonov's Theorem (see, e.g., [AW00]), this implies that f is an L^2 coboundary: there exists $u \in L^2$ such that $f = u - u \circ T$ almost everywhere. Then

$$\hat{T}_t(e^{-\mathbf{i}tu}) = \hat{T}(e^{\mathbf{i}tf} e^{-\mathbf{i}tu}) = \hat{T}(e^{-\mathbf{i}tu \circ T}) = e^{-\mathbf{i}tu}. \tag{3.72}$$

By [ITM50], this yields $e^{-\mathbf{i}tu} \in \mathcal{L}$. In particular, the function $e^{-\mathbf{i}tu}$ is continuous for any small enough t. Lemma 3.21 shows that u itself is continuous. In particular, there exists a cylinder $[b_0, \dots, b_k]$ on which u is bounded. Since f is bounded on each element of the partition α, the equation $f = u - u \circ T$ implies that u is bounded on b_k. Together with the topological transitivity of T, we obtain that u is bounded on each $a \in \alpha$.

Let $\{a_1, \dots, a_n\}$ be a finite subset of α such that each element of α contains the image of one of the a_is (it exists by the big preimage property). Let $a \in \alpha$, and choose i such that $a \subset T(a_i)$; then the equation $f = u - u \circ T$ gives $\|u 1_a\|_{L^\infty} \leq \|(|f| + |u|) 1_{a_i}\|_{L^\infty}$. This shows that u is uniformly bounded, as desired. ∎

In fact, a slightly refined version of the same argument also shows that u is Hölder continuous.

LEMMA 3.21: *Let u be a real function on a metric space X, and assume that $e^{\mathbf{i}tu}$ is continuous for $t \in [a,b]$ a nontrivial interval of $\mathbb{R}$. Then u is continuous.*

Proof. We will show that, if v_n is a real sequence such that $e^{\mathbf{i}tv_n}$ converges to 0 for any $t \in [a,b]$, then $v_n \to 0$. Applying this result to $v_n = u(x_n) - u(x)$ when $x_n \to x$, this gives the required continuity of u at $x \in X$, for any x.

Let $A_N = \{t \in [a,b] \mid \forall n \geq N, \mathrm{dist}(tv_n, 2\pi\mathbb{Z}) \leq 1\}$. The set A_N is a closed subset of $[a,b]$, and $\bigcup A_N = [a,b]$. By Baire's Theorem, there exists a set A_N containing a nontrivial interval $[c,d]$. For $n \geq N$ and $t \in [c,d]$, the number tv_n belongs to $2\pi\mathbb{Z} + [-1,1]$, and depends continuously on t. It has to stay in the same connected component of $2\pi\mathbb{Z} + [-1,1]$, therefore $|cv_n - dv_n| \leq 2$. This shows that v_n is bounded.

Any cluster value v of v_n satisfies $e^{\mathbf{i}tv} = 0$ for any $t \in [a,b]$, hence $v = 0$. ∎

Appendix A. The Berry–Esseen theorem for Gibbs–Markov maps

In this appendix, we obtain necessary and sufficient conditions for the Berry–Esseen theorem, for Gibbs–Markov maps.

THEOREM A.1: *Let $T : X \to X$ be a probability preserving mixing Gibbs–Markov map, and let $f : X \to \mathbb{R}$ satisfy $\sum_{a\in\alpha} m(a) Df(a)^\eta < \infty$ for some $\eta \in (0,1]$. Assume $f \in L^2$ and $E(f) = 0$, and $S_n f/\sqrt{n} \to \mathcal{N}(0,\sigma^2)$ with $\sigma^2 > 0$. Let*

$$\Delta_n := \sup_{x\in\mathbb{R}} \left| m\{S_n f/\sqrt{n} < x\} - P(\mathcal{N}(0,\sigma^2) < x)\right|. \tag{A.1}$$

Let $\delta \in (0,1)$. Then $\Delta_n = O(n^{-\delta/2})$ if and only if $E(f^2 1_{|f|>x}) = O(x^{-\delta})$ when $x \to \infty$.

Moreover, $\Delta_n = O(n^{-1/2})$ if and only if $E(f^2 1_{|f|>x}) = O(x^{-1})$ when $x \to \infty$, and $E(f^3 1_{|f|<x})$ is uniformly bounded.

When one considers i.i.d. random variables instead of Birkhoff sums, this theorem for $\delta < 1$ is proved in [IL71, Theorem 3.4.1], and the proof for $\delta = 1$ is given in [Ibr66]. For the proof in the dynamical setting, we will essentially follow the same strategy as in the i.i.d. case, the additional crucial ingredient being the estimate on $\lambda(t)$ provided by Theorem 3.13. We will only give the proof for $\delta < 1$, since the proof for $\delta = 1$ is very similar following the arguments of [Ibr66].

Proof of the necessity in Theorem A.1. Assuming $\Delta_n = O(n^{-\delta/2})$, we will prove $E(f^2 1_{|f|>x}) = O(x^{-\delta})$. This is trivial if $f \in L^3$, so we can assume this is not the case. In this proof, ϵ will denote the minimum of $\epsilon(p)$ given by Theorem 3.13 for $p \in [2,3]$. Consider $p \in [2,3]$ such that $f \in L^p$ and $f \notin L^{p+\epsilon/2}$, and let $q = \min(p+\epsilon, 3)$. Hence, $\lambda(t) = E(e^{\mathbf{i}tf}) + ct^2 + O(|t|^q)$ for some $c \in \mathbb{R}$. It will be more convenient to write this estimate as follows:

$$\lambda(t) = E(e^{\mathbf{i}tf})e^{ct^2+t^2\phi(t)} \text{ with } \phi(t) = O(|t|^{q-2}). \tag{A.2}$$

Let W be the symmetrization of f, i.e., the difference of two independent copies of f. Its characteristic function is $E(e^{\mathbf{i}tW}) = |E(e^{\mathbf{i}tf})|^2$. Let us write $E(e^{\mathbf{i}tW}) = e^{-\sigma_0^2 t^2 + t^2\gamma_0(t)}$ where $\sigma_0^2 = E(f^2)$ and γ_0 is a real function defined on a neighborhood of 0. [IL71, Paragraph III.4] proves the following fact:

$$\begin{aligned} &\text{If } \int_0^x t^2|\gamma_0(t)| = O(x^{3+\tilde\delta}), 0<\tilde\delta<1, \text{ when } x\to 0,\\ &\qquad\qquad \text{then } E(f^2 1_{|f|>x}) = O(x^{-\tilde\delta}) \text{ when } x \to +\infty. \end{aligned} \tag{A.3}$$

To conclude, it is therefore sufficient to estimate $\int t^2|\gamma_0(t)|$.

Let H denote the distribution function of $\mathcal{N}(0, 2\sigma^2)$, and F_n the distribution function of the difference of two independent copies of $S_n f/\sqrt{n}$. From the assumption $\Delta_n = O(n^{-\delta/2})$, it follows that $\sup_{x\in\mathbb{R}} |H(x) - F_n(x)| \leq Cn^{-\delta/2}$. Let $h(t)$ and $f_n(t)$ be the characteristic functions of H and F_n, i.e., $h(t) = e^{-\sigma^2t^2}$ and $f_n(t) = |E(e^{\mathbf{i}tS_n f/\sqrt{n}})|^2$. Integrating by parts the equality $f_n(t) - h(t) = \int e^{\mathbf{i}tx}\,\mathrm{d}(F_n(x) - H(x))$, we obtain

$$\frac{f_n(t) - h(t)}{\mathbf{i}t} = \int e^{\mathbf{i}tx}(F_n(x) - H(x))\,\mathrm{d}x. \tag{A.4}$$

This shows that the L^2 functions $(f_n(t) - h(t))/\mathbf{i}t$ and $F_n - H$ are Fourier transforms of one another. The functions $te^{-t^2/2}$ and $-\mathbf{i}xe^{-x^2/2}/\sqrt{2\pi}$ are also Fourier transforms of one another. Hence, Parseval's theorem gives

$$\int \frac{f_n(t) - h(t)}{t} \cdot te^{-t^2/2}\,\mathrm{d}t = C\int (F_n(x) - H(x))xe^{-x^2/2} = O(n^{-\delta/2}). \tag{A.5}$$

Since $\int_{|t|\geq \log n} e^{-t^2/2}\,\mathrm{d}t = O(n^{-\delta/2})$, this yields

$$\int_{|t|\leq \log n} (f_n(t) - h(t))e^{-t^2/2}\,\mathrm{d}t = O(n^{-\delta/2}). \tag{A.6}$$

The characteristic expansion of f gives

$$f_n(t) = |E(e^{\mathbf{i}tS_n f/\sqrt{n}})|^2 = \left|\lambda\left(\frac{t}{\sqrt{n}}\right)\right|^{2n} \left|\mu\left(\frac{t}{\sqrt{n}}\right)\right|^2 + \epsilon_n(t),$$

where $\epsilon_n(t)$ tends exponentially fast to 0 (by Theorem 3.1), and the function μ satisfies $\mu(t) = 1 + O(t)$ (by Remark 3.18). Let $g_n(t) = |\lambda(t/\sqrt{n})\ |^{2n}$; then $\int_{|t|\leq \log n}(f_n(t) - g_n(t))e^{-t^2/2}\,\mathrm{d}t = O(n^{-1/2})$. Therefore, (A.6) gives

$$\text{(A.7)} \qquad \int_{|t|\leq \log n}(g_n(t) - h(t))e^{-t^2/2}\,\mathrm{d}t = O(n^{-\delta/2}).$$

Moreover, by (A.2)

$$\begin{aligned} g_n(t) &= |E(e^{\mathbf{i}tf/\sqrt{n}})|^{2n} e^{2ct^2+2t^2\,\mathrm{Re}\,\phi(t/\sqrt{n})} \\ &= e^{-\sigma_0^2 t^2 + t^2\gamma_0(t/\sqrt{n})+2ct^2+2t^2\,\mathrm{Re}\,\phi(t/\sqrt{n})} \\ &= e^{-\sigma^2 t^2 + t^2\gamma_0(t/\sqrt{n})+2t^2\,\mathrm{Re}\,\phi(t/\sqrt{n})}. \end{aligned}$$

Let $h_n(t) = e^{-\sigma^2 t^2 + t^2\gamma_0(t/\sqrt{n})}$. Since $\phi(t) = O(|t|^{q-2})$ by (A.2), we have $\int_{|t|\leq \log n}(g_n(t) - h_n(t))e^{-t^2/2}\,\mathrm{d}t = O(n^{-(q-2)/2})$. Hence,

$$\text{(A.8)} \qquad \int_{|t|\leq \log n}(h_n(t) - h(t))e^{-t^2/2}\,\mathrm{d}t = O(n^{-\delta/2}) + O(n^{-(q-2)/2}).$$

Since $h_n(t) - h(t) = e^{-\sigma^2 t^2}(e^{t^2\gamma_0(t/\sqrt{n})} - 1)$, we can now conclude as in [IL71, Page 106] to get

$$\text{(A.9)} \qquad \int_0^x t^2|\gamma_0(t)| = O(x^{3+\delta}) + O(x^{q+1}).$$

By (A.3), this proves that $E(f^2 1_{|f|>x}) = O(x^{-\tilde{\delta}})$ for $\tilde{\delta} = \min(\delta, q-2)$. If $q - 2 < \delta$ (in particular, $q \neq 3$, so $q = p + \epsilon$), we have $\tilde{\delta} = q - 2$, hence f belongs to $L^{q'}$ for any $q' < q$. In particular, $f \in L^{q-\epsilon/2} = L^{p+\epsilon/2}$. This is not compatible with the choice of p. Hence, $q - 2 \geq \delta$, whence $\tilde{\delta} = \delta$, and $E(f^2 1_{|f|>x}) = O(x^{-\delta})$. ∎

Proof of the sufficiency in Theorem A.1. Assuming $E(f^2 1_{|f|>x}) = O(x^{-\delta})$, we will prove $\Delta_n = O(n^{-\delta/2})$. We essentially follow the arguments of the proof of the necessity, in the reverse direction, the main difference being that we no longer need to work with the symmetrization of the random variables.

Let us write $E(e^{\mathbf{i}tf}) = e^{-\sigma_0^2 t^2/2 + t^2\gamma(t)}$. [IL71, Page 111] proves that, under the assumption $E(f^2 1_{|f|>x}) = O(x^{-\delta})$, the function γ satisfies $\int_0^x t^2|\gamma(t)|\,\mathrm{d}t = O(x^{3+\delta})$.

Moreover, f belongs to L^p for any $p < 2+\delta$. Let $q = \min(2+\delta+\epsilon/2, 3) > 2+\delta$. Taking $p = 2+\delta-\epsilon/2$, Theorem 3.13 shows that $\lambda(t) = E(e^{\mathbf{i}tf}) + ct^2 + O(|t|^q)$, which we may rewrite as $\lambda(t) = E(e^{\mathbf{i}tf})e^{ct^2+t^2\phi(t)}$ where $\phi(t) = O(|t|^{q-2})$. Together with the expansion of $E(e^{\mathbf{i}tf})$, we obtain

$$\lambda(t) = e^{-\sigma^2 t^2/2 + t^2\psi(t)} \text{ with } \int_0^x t^2|\psi(t)| \,\mathrm{d}t = O(x^{3+\delta}). \tag{A.10}$$

Let f_n denote the characteristic function of $S_n f/\sqrt{n}$. The classical Berry–Esseen estimate [IL71, Theorem 1.5.2] shows that, for any $T > 0$,

$$\Delta_n \leq C \int_{-T}^{T} \frac{1}{|t|} |f_n(t) - e^{-\sigma^2 t^2/2}| \,\mathrm{d}t + C/T. \tag{A.11}$$

Let us choose $T = \rho\sqrt{n}$ with ρ small enough. The second term in this estimate is then $O(n^{-1/2}) = O(n^{-\delta})$. For the first term, we split the integral in two parts, corresponding to $|t| \leq 1/n$ and $|t| > 1/n$. In the first part, we have

$$|f_n(t) - 1| = \left| E(e^{\mathbf{i}tS_n f/\sqrt{n}} - 1) \right| \leq |t| E|S_n f|/\sqrt{n} \leq \sqrt{n}|t|. \tag{A.12}$$

The resulting integral is bounded by

$$\int_{|t|\leq 1/n} \sqrt{n} + |t|^{-1}|1 - e^{-\sigma^2 t^2/2}| \,\mathrm{d}t = O(n^{-1/2}). \tag{A.13}$$

Hence,

$$\Delta_n \leq C \int_{1/n \leq |t| \leq \rho\sqrt{n}} \frac{1}{|t|} |f_n(t) - e^{-\sigma^2 t^2/2}| \,\mathrm{d}t + O(n^{-1/2}). \tag{A.14}$$

We have $f_n(t) = \lambda(t/\sqrt{n})^n \mu(t/\sqrt{n}) + \epsilon_n(t)$, where $\epsilon_n(t)$ tends exponentially fast to 0, while $\mu(t) = 1 + O(t)$. Let $g_n(t) = \lambda(t/\sqrt{n})^n$. By (A.10), if ρ is small enough, we have $|\lambda(t)| \leq e^{-\sigma^2 t^2/4}$ for $|t| \leq \rho$. This yields $|\lambda(t/\sqrt{n})|^n \leq e^{-\sigma^2 t^2/4}$ for $|t| \leq \rho\sqrt{n}$. Hence,

$$\int_{1/n \leq |t| \leq \rho\sqrt{n}} \frac{1}{|t|} |f_n(t) - g_n(t)| \leq C/\sqrt{n}. \tag{A.15}$$

With (A.10), we obtain

$$\begin{aligned}
\Delta_n &\leq C \int_{1/n \leq |t| \leq \rho\sqrt{n}} \frac{1}{|t|} |g_n(t) - e^{-\sigma^2 t^2/2}| \,\mathrm{d}t + O(n^{-1/2}) \\
&= C \int_{1/n \leq |t| \leq \rho\sqrt{n}} \frac{1}{|t|} e^{-\sigma^2 t^2/2} |e^{t^2\psi(t/\sqrt{n})} - 1| \,\mathrm{d}t + O(n^{-1/2}).
\end{aligned}$$

Since $\int_0^x t^2|\psi(t)| \,\mathrm{d}t = O(x^{3+\delta})$, this last integral is bounded by $O(n^{-\delta/2})$ (see, e.g., [IL71, bottom of Page 107]). This concludes the proof. ∎

References

[AD98] J. Aaronson and M. Denker, *Characteristic functions of random variables attracted to 1-stable laws*, The Annals of Probability **26** (1998), 399–415.

[AD01a] J. Aaronson and M. Denker, *A local limit theorem for stationary processes in the domain of attraction of a normal distribution*, in *Asymptotic Methods in Probability and Statistics with Applications (St. Petersburg, 1998)*, Stat. Ind. Technol., Birkhäuser Boston, Boston, MA, 2001, pp. 215–223.

[AD01b] J. Aaronson and M. Denker, *Local limit theorems for partial sums of stationary sequences generated by Gibbs–Markov maps*, Stochastics and Dynamics **1** (2001), 193–237.

[AW00] J. Aaronson and B. Weiss, *Remarks on the tightness of cocycles*, Colloquium Mathematicum **84/85** (2000), 363–376, Dedicated to the memory of Anzelm Iwanik.

[BGT87] N. H. Bingham, C. M. Goldie and J. L. Teugels, *Regular variation*, in *Encyclopedia of Mathematics and its Applications*, Vol. 27, Cambridge University Press, Cambridge, 1987.

[Bil95] P. Billingsley, *Probability and Measure*, third edn., Wiley Series in Probability and Mathematical Statistics, Wiley, New York, 1995.

[BL76] J. Bergh and J. Löfström, *Interpolation Spaces. An Introduction*, Springer-Verlag, Berlin, 1976, Grundlehren der Mathematischen Wissenschaften, No. 223.

[Cam64] S. Campanato, *Proprietà di una famiglia di spazi funzionali*, Annali della Scuola Normale Superiore di Pisa (3) **18** (1964), 137–160.

[DJ89] M. Denker and A. Jakubowski, *Stable limit distributions for strongly mixing sequences*, Statistics & Probability Letters **8** (1989), 477–483.

[Fel66] W. Feller, *An Introduction to Probability Theory and its Applications. Vol. II*, Wiley, New York, 1966.

[GH88] Y. Guivarc'h and J. Hardy, *Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d'Anosov*, Annales de l'Institut Henri Poincaré. Probabilités et Statistiques **24** (1988), 73–98.

[GL06] S. Gouëzel and C. Liverani, *Banach spaces adapted to Anosov systems*, Ergodic Theory and Dynamical Systems **26** (2006), 189–217.

[Gou04] S. Gouëzel, *Central limit theorem and stable laws for intermittent maps*, Probability Theory and Related Fields **128** (2004), 82–122.

[Hen93] H. Hennion, *Sur un théorème spectral et son application aux noyaux lipchitziens*, Proceedings of the American Mathematical Society **118** (1993), 627–634.

[Her05] L. Hervé, *Théorème local pour chaînes de Markov de probabilité de transition quasi-compacte. Applications aux chaînes V-géométriquement ergodiques et aux modèles itératifs*, Annales de l'Institut Henri Poincaré. Probabilités et Statistiques **41** (2005), 179–196.

[HP08] L. Hervé and F. Pène, *Nagaev method via Keller–Liverani theorem*, Preprint, 2008.

[Ibr66] I. A. Ibragimov, *On the accuracy of approximation by the normal distribution of distribution functions of sums of independent random variables*, Teoriya Veroyatnosteĭ i ee Primeneniya **11** (1966), 632–655.

[IL71] I. A. Ibragimov and Y. V. Linnik, *Independent and stationary sequences of random variables*, Wolters-Noordhoff Publishing, Groningen, 1971, With a supplementary chapter by I. A. Ibragimov and V. V. Petrov, Translation from the Russian edited by J. F. C. Kingman.

[ITM50] C. T. Ionescu Tulcea and G. Marinescu, *Théorie ergodique pour des classes d'opérations non complètement continues*, Annals of Mathematics (2) **52** (1950), 140–147.

[Jak93] A. Jakubowski, *Minimal conditions in p-stable limit theorems*, Stochastic Processes and their Applications **44** (1993), 291–327.

[Kat66] T. Kato, *Perturbation Theory for Linear Operators*, Die Grundlehren der mathematischen Wissenschaften, Band 132, Springer-Verlag, New York, 1966.

[KL99] G. Keller and C. Liverani, *Stability of the spectrum for transfer operators*, Annali della Scuola Normale Superiore di Pisa (4) **28** (1999), 141–152.

[RE83] J. Rousseau-Egele, *Un théorème de la limite locale pour une classe de transformations dilatantes et monotones par morceaux*, Annals of Probability **11** (1983), 772–788.

[Sar06] O. Sarig, *Continuous phase transitions for dynamical systems*, Communications in Mathematical Physics **267** (2006), 631–667.